SECOND DECLARATION OF PAUL POLAKIS, Ph.D.

I, Paul Polakis, Ph.D., declare and say as follows:

1. I am currently employed by Genentech, Inc. where my job title is Staff Scientist.

2. Since joining Genentech in 1999, one of my primary responsibilities has been leading Genentech's Tumor Antigen Project, which is a large research project with a primary focus on identifying tumor cell markers that find use as targets for both the diagnosis and treatment of cancer in humans.

3. As I stated in my previous Declaration dated May 7, 2004 (attached as Exhibit A), my laboratory has been employing a variety of techniques, including microarray analysis, to identify genes which are differentially expressed in human tumor tissue relative to normal human tissue. The primary purpose of this research is to identify proteins that are abundantly expressed on certain human tumor tissue(s) and that are either (i) not expressed, or (ii) expressed at detectably lower levels, on normal tissue(s).

4. In the course of our research using microarray analysis, we have identified approximately 200 gene transcripts that are present in human tumor tissue at significantly higher levels than in normal human tissue. To date, we have successfully generated antibodies that bind to 31 of the tumor antigen proteins expressed from these differentially expressed gene transcripts and have used these antibodies to quantitatively determine the level of production of these tumor antigen proteins in both human tumor tissue and normal tissue. We have then quantitatively compared the levels of mRNA and protein in both the tumor and normal tissues analyzed. The results of these analyses are attached herewith as Exhibit B. In Exhibit B, "+" means that the mRNA or protein was detectably overexpressed in the tumor tissue relative to normal tissue and "-" means that no detectable overexpression was observed in the tumor tissue relative to normal tissue.

5. As shown in Exhibit B, of the 31 genes identified as being detectably overexpressed in human tumor tissue as compared to normal human tissue at the mRNA level, 28 of them (i.e., greater than 90%) are also detectably overexpressed in human tumor tissue as compared to normal human tissue at the protein level. As such, in the cases where we have been able to quantitatively measure both (i) mRNA and (ii) protein levels in both (i) tumor tissue and (ii) normal tissue, we have observed that in the vast majority of cases, there is a very strong correlation between increases in mRNA expression and increases in the level of protein encoded by that mRNA.

6. Based upon my own experience accumulated in more than 20 years of research, including the data discussed in paragraphs 4-5 above and my knowledge of the relevant scientific literature, it is my considered scientific opinion that for human genes, an increased level of mRNA in a tumor tissue relative to a normal tissue more often than not correlates to a similar increase in abundance of the encoded protein in the tumor tissue relative to the normal tissue. In fact, it remains a generally accepted working assumption in molecular biology that increased mRNA levels are more often than not predictive of elevated levels of the encoded protein. In fact, an entire industry focusing on the research and development of therapeutic antibodies to treat a variety of human diseases, such as cancer, operates on this working assumption.

7. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information or belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful statements may jeopardize the validity of the application or any patent issued thereon.

Dated: 3-29-06

By: Paul Polakis

Paul Polakis, Ph.D.

EXHIBIT A

DECLARATION OF PAUL POLAKIS, Ph.D.

I, Paul Polakis, Ph.D., declare and say as follows:

1. I was awarded a Ph.D. by the Department of Biochemistry of the Michigan State University in 1984. My scientific Curriculum Vitae is attached to and forms part of this Declaration (Exhibit A).

2. I am currently employed by Genentech, Inc. where my job title is Staff Scientist. Since joining Genentech in 1999, one of my primary responsibilities has been leading Genentech's Tumor Antigen Project, which is a large research project with a primary focus on identifying tumor cell markers that find use as targets for both the diagnosis and treatment of cancer in humans.

3. As part of the Tumor Antigen Project, my laboratory has been analyzing differential expression of various genes in tumor cells relative to normal cells. The purpose of this research is to identify proteins that are abundantly expressed on certain tumor cells and that are either (i) not expressed, or (ii) expressed at lower levels, on corresponding normal cells. We call such differentially expressed proteins "tumor antigen proteins". When such a tumor antigen protein is identified, one can produce an antibody that recognizes and binds to that protein. Such an antibody finds use in the diagnosis of human cancer and may ultimately serve as an effective therapeutic in the treatment of human cancer.

4. In the course of the research conducted by Genentech's Tumor Antigen Project, we have employed a variety of scientific techniques for detecting and studying differential gene expression in human tumor cells relative to normal cells, at genomic DNA, mRNA and protein levels. An important example of one such technique is the well known and widely used technique of microarray analysis which has proven to be extremely useful for the identification of mRNA molecules that are differentially expressed in one tissue or cell type relative to another. In the course of our research using microarray analysis, we have identified approximately 200 gene transcripts that are present in human tumor cells at significantly higher levels than in corresponding normal human cells. To date, we have generated antibodies that bind to about 30 of the tumor antigen proteins expressed from these differentially expressed gene transcripts and have used these antibodies to quantitatively determine the level of production of these tumor antigen proteins in both human cancer cells and corresponding normal cells. We have then compared the levels of mRNA and protein in both the tumor and normal cells analyzed.

5. From the mRNA and protein expression analyses described in paragraph 4 above, we have observed that there is a strong correlation between changes in the level of mRNA present in any particular cell type and the level of protein

expressed from that mRNA in that cell type. In approximately 80% of our observations we have found that increases in the level of a particular mRNA correlates with changes in the level of protein expressed from that mRNA when human tumor cells are compared with their corresponding normal cells.

6. Based upon my own experience accumulated in more than 20 years of research, including the data discussed in paragraphs 4 and 5 above and my knowledge of the relevant scientific literature, it is my considered scientific opinion that for human genes, an increased level of mRNA in a tumor cell relative to a normal cell typically correlates to a similar increase in abundance of the encoded protein in the tumor cell relative to the normal cell. In fact, it remains a central dogma in molecular biology that increased mRNA levels are predictive of corresponding increased levels of the encoded protein. While there have been published reports of genes for which such a correlation does not exist, it is my opinion that such reports are exceptions to the commonly understood general rule that increased mRNA levels are predictive of corresponding increased levels of the encoded protein.

7. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information or belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful statements may jeopardize the validity of the application or any patent issued thereon.

Dated: 5/07/04

By: Paul Polakis

Paul Polakis, Ph.D.

SV 2031808 v1

CURRICULUM VITAE

PAUL G. POLAKIS
Staff Scientist
Genentech, Inc
1 DNA Way, MS#40
S. San Francisco, CA 94080

EDUCATION:

Ph.D., Biochemistry, Department of Biochemistry,
Michigan State University (1984)

B.S., Biology. College of Natural Science, Michigan State University (1977)

PROFESSIONAL EXPERIENCE:

2002-present	Staff Scientist, Genentech, Inc S. San Francisco, CA
1999- 2002	Senior Scientist, Genentech, Inc., S. San Francisco, CA
1997 -1999	Research Director Onyx Pharmaceuticals, Richmond, CA
1992- 1996	Senior Scientist, Project Leader, Onyx Pharmaceuticals, Richmond, CA
1991-1992	Senior Scientist, Chiron Corporation, Emeryville, CA.
1989-1991	Scientist, Cetus Corporation, Emeryville CA.
1987-1989	Postdoctoral Research Associate, Genentech, Inc., South SanFrancisco, CA.
1985-1987	Postdoctoral Research Associate, Department of Medicine, Duke University Medical Center, Durham, NC

1984-1985	Assistant Professor, Depar[illegible]ent of Chemistry, Oberlin College, Oberlin, O[illegible]
1980-1984	Graduate Research Assistant, Department of Biochemistry, Michigan State University East Lansing, Michigan

PUBLICATIONS:

1. **Polakis, P G.** and Wilson, J. E. 1982 Purification of a Highly Bindable Rat Brain Hexokinase by High Performance Liquid Chromatography. **Biochem. Biophys. Res. Commun.** 107, 937-943.

2. **Polakis, P.G.** and Wilson, J. E. 1984 Proteolytic Dissection of Rat Brain Hexokinase: Determination of the Cleavage Pattern during Limited Digestion with Trypsin. **Arch. Biochem. Biophys.** 234, 341-352.

3. **Polakis, P. G.** and Wilson, J. E. 1985 An Intact Hydrophobic N-Terminal Sequence is Required for the Binding Rat Brain Hexokinase to Mitochondria. **Arch. Biochem. Biophys.** 236, 328-337.

4. Uhing, R.J., **Polakis,P.G.** and Snyderman, R. 1987 Isolaton of GTP-binding Proteins from Myeloid HL60 Cells. **J. Biol. Chem.** 262, 15575-15579.

5. **Polakis, P.G.**, Uhing, R.J. and Snyderman, R. 1988 The Formylpeptide Chemoattractant Receptor Copurifies with a GTP-binding Protein Containing a Distinct 40 kDa Pertussis Toxin Substrate. **J. Biol. Chem.** 263, 4969-4979.

6. Uhing, R. J., Dillon, S., **Polakis, P. G.**, Truett, A. P. and Snyderman, R. 1988 Chemoattractant Receptors and Signal Transduction Processes in Cellular and Molecular Aspects of Inflammation (Poste, G. and Crooke, S. T. eds.) pp 335-379.

7. **Polakis, P.G.**, Evans, T. and Snyderman 1989 Multiple Chromatographic Forms of the Formylpeptide Chemoattractant Receptor and their Relationship to GTP-binding Proteins. **Biochem. Biophys. Res. Commun.** 161, 276-283.

8. **Polakis, P. G.**, Snyderman, R. and Evans, T. 1989 Characterization of G25K, a GTP-binding Protein Containing a Novel Putative Nucleotide Binding Domain. **Biochem. Biophys. Res. Comun.** 160, 25-32.

9. **Polakis, P.**, Weber,R.F., Nevins,B., Didsbury, J. Evans,T. and Snyderman, R. 1989 Identification of the ral and rac1 Gene Products, Low Molecular Mass GTP-binding Proteins from Human Platelets. **J. Biol. Chem.** 264, 16383-16389.

10. Snyderman, R., Perianin, A., Evans, T., **Polakis, P.** and Didsbury, J. 1989 G Proteins and Neutrophil Function. In ADP-Ribosylating Toxins and G Proteins: Insights into Signal Transduction. (J. Moss and M. Vaughn, eds.) Amer. Soc. Microbiol. pp. 295-323.

11. Hart, M.J., **Polakis, P.G.** Evans, T. and Cerrione, R.A. 1990 The Identification and Charaterization of an Epidermal Growth Factor-Stimulated Phosphorylation of a Specific Low Molecular Mass GTP-binding Protein in a Reconstituted Phospholipid Vesicle System. **J. Biol. Chem.** 265, 5990-6001.

12. Yatani, A., Okabe, K., **Polakis, P.** Halenbeck, R. McCormick, F. and Brown, A. M. 1990 ras p21 and GAP Inhibit Coupling of Muscarinic Receptors to Atrial K^+ Channels. **Cell.** 61, 769-776.

13. Munemitsu, S., Innis, M.A., Clark, R., McCormick, F., Ullrich, A. and **Polakis, P.G.** 1990 Molecular Cloning and Expression of a G25K cDNA, the Human Homolog of the Yeast Cell Cycle Gene CDC42. **Mol. Cell. Biol.** 10, 5977-5982.

14. **Polakis, P.G.** Rubinfeld, B. Evans, T. and McCormick, F. 1991 Purification of Plasma Membrane-Associated GTPase Activating Protein (GAP) Specific for rap-1/krev-1 from HL60 Cells. **Proc. Natl. Acad. Sci. USA** 88, 239-243.

15. Moran, M. F., **Polakis, P.**, McCormick, F., Pawson, T. and Ellis, C. 1991 Protein Tyrosine Kinases Regulate the Phosphorylation, Protein Interactions, Subcellular Distribution, and Activity of p21ras GTPase Activating Protein. **Mol. Cell. Biol.** 11, 1804-1812

16. Rubinfeld, B., Wong, G., Bekesi, E. Wood, A. McCormick, F. and **Polakis, P. G.** 1991 A Synthetic Peptide Corresponding to a Sequence in the GTPase Activating Protein Inhibits $p21^{ras}$ Stimulation and Promotes Guanine Nucleotide Exchange. **Internatl. J. Peptide and Prot. Res.** 38, 47-53.

17. Rubinfeld, B., Munemitsu, S., Clark, R., Conroy, L., Watt, K., Crosier, W., McCormick, F., and **Polakis, P.** 1991 Molecular Cloning of a GTPase Activating Protein Specific for the Krev-1 Protein $p21^{rap1}$. **Cell** 65, 1033-1042.

18. Zhang, K. Papageorge, A., G., Martin, P., Vass, W. C., Olah, Z., **Polakis, P.**, McCormick, F. and Lowy, D, R. 1991 Heterogenous Amino Acids in RAS and Rap1A Specifying Sensitivity to GAP Proteins. **Science** 254, 1630-1634.

19. Martin, G., Yatani, A., Clark, R., **Polakis, P.**, Brown, A. M. and McCormick, F. 1992 GAP Domains Responsible for $p21^{ras}$-dependent Inhibition of Muscarinic Atrial K^+ Channel Currents. **Science** 255, 192-194.

20. McCormick, F., Martin, G. A., Clark, R., Bollag, G. and **Polakis, P** . 1992 Regulation of p21ras by GTPase Activating Proteins. Cold Spring Harbor **Symposia on Quantitative Biology**. Vol. 56, 237-241.

21. Pronk, G. B., **Polakis, P.**, Wong, G., deVries-Smits, A. M., Bos J. L. and McCormick, F. 1992 $p60^{v\text{-}src}$ Can Associate with and Phosphorylate the $p21^{ras}$ GTPase Activating Protein. **Oncogene** 7,389-394.

22. **Polakis P.** and McCormick, F. 1992 Interactions Between $p21^{ras}$ Proteins and Their GTPase Activating Proteins. In **Cancer Surveys** (Franks, L. M., ed.) 12, 25-42.

23. Wong, G., Muller, O., Clark, R., Conroy, L., Moran, M., **Polakis, P.** and McCormick, F. 1992 Molecular coloning and nucleic acid binding properties of the GAP-associated tyrosine phosphoprotein p62. **Cell** 69, 551-558.

24. **Polakis, P.**, Rubinfeld, B. and McCormick, F. 1992 Phosphorylation of rap1GAP in vivo and by cAMP-dependent Kinase and the Cell Cycle $p34^{cdc2}$ Kinase in vitro. **J. Biol. Chem.** 267, 10780-10785.

25. McCabe, P.C., Haubrauck, H., **Polakis, P.**, McCormick, F., and Innis, M. A. 1992 Functional Interactions Between $p21^{rap1A}$ and Components of the Budding pathway of *Saccharomyces cerevisiae*. **Mol. Cell. Biol.** 12, 4084-4092.

26. Rubinfeld, B., Crosier, W.J., Albert, I., Conroy,L., Clark, R., McCormick, F. and **Polakis, P.** 1992 Localization of the rap1GAP Catalytic Domain and Sites of Phosphorylation by Mutational Analysis. **Mol. Cell . Biol.** 12, 4634-4642.

27. Ando, S., Kaibuchi, K., Sasaki, K., Hiraoka, T., Nishiyama, T., Mizuno, T., Asada, M., Nunoi, H., Matsuda, I., Matsuura, Y., **Polakis, P.**, McCormick, F. and Takai, Y. 1992 Post-translational processing of rac p21s is important both for their interaction with the GDP/GTP exchange proteins and for their activation of NADPH oxidase. **J. Biol. Chem.** 267, 25709-25713.

28. Janoueix-Lerosey, I., **Polakis, P.**, Tavitian, A. and deGunzberg, J. 1992 Regulation of the GTPase activity of the ras-related rap2 protein. **Biochem. Biophys. Res. Commun.** 189, 455-464.

29. **Polakis, P.** 1993 GAPs Specific for the rap1/Krev-1 Protein. in GTP-binding Proteins: the *ras*-superfamily. (J.C. LaCale and F. McCormick, eds.) 445-452.

30. **Polakis, P.** and McCormick, F. 1993 Structural requirements for the interaction of $p21^{ras}$ with GAP, exchange factors, and its bological effector target. **J. Biol Chem.** 268, 9157-9160.

31. Rubinfeld, B., Souza, B. Albert, I., Muller, O., Chamberlain, S., Masiarz, F., Munemitsu, S. and **Polakis, P.** 1993 Association of the APC gene product with beta- catenin. **Science** 262, 1731-1734.

32. Weiss, J., Rubinfeld, B., **Polakis, P.**, McCormick, F. Cavenee, W. A. and Arden, K. 1993 The gene for human rap1-GTPase activating protein (rap1GAP) maps to chromosome 1p35-1p36.1. **Cytogenet. Cell Genet.** 66, 18-21.

33. Sato, K. Y., **Polakis, P.**, Haubruck, H., Fasching, C. L., McCormick, F. and Stanbridge, E. J. 1994 Analysis of the tumor suppressor acitvity of the K-rev gene in human tumor cell lines. **Cancer Res.** 54, 552-559.

34. Janoueix-Lerosey, I., Fontenay, M., Tobelem, G., Tavitian, A., **Polakis, P.** and DeGunzburg, J. 1994 Phosphorylation of rap1GAP during the cell cycle. **Biochem. Biophys. Res. Commun.** 202, 967-975

35. Munemitsu, S., Souza, B., Mueller, O., Albert, I., Rubinfeld, B., and **Polakis, P.** 1994 The APC gene product associates with microtubules in vivo and affects their assembly in vitro. **Cancer Res.** 54, 3676-3681.

36. Rubinfeld, B. and **Polakis, P.** 1995 Purification of baculovirus produced rap1GAP. **Methods Enz.** 255,31

37. **Polakis, P.** 1995 Mutations in the APC gene and their implications for protein structure and function. **Current Opinions in Genetics and Development** 5, 66-71

38. Rubinfeld, B., Souza, B., Albert, I., Munemitsu, S. and **Polakis P.** 1995 The APC protein and E-cadherin form similar but independent complexes with α-catenin, β-catenin and Plakoglobin. **J. Biol. Chem.** 270, 5549-5555

39. Munemitsu, S., Albert, I., Souza, B., Rubinfeld, B., and **Polakis, P.** 1995 Regulation of intracellular β-catenin levels by the APC tumor suppressor gene. **Proc. Natl. Acad. Sci.** 92, 3046-3050.

40. Lock, P., Fumagalli, S., **Polakis, P.** McCormick, F. and Courtneidge, S. A. 1996 The human p62 cDNA encodes Sam68 and not the rasGAP-associated p62 protein. **Cell** 84, 23-24.

41. Papkoff, J., Rubinfeld, B., Schryver, B. and **Polakis, P.** 1996 Wnt-1 regulates free pools of catenins and stabilizes APC-catenin complexes. **Mol. Cell. Biol.** 16, 2128-2134.

42. Rubinfeld, B., Albert, I., Porfiri, E., Fiol, C., Munemitsu, S. and **Polakis, P.** 1996 Binding of GSK3β to the APC-β-catenin complex and regulation of complex assembly. **Science** 272, 1023-1026.

43. Munemitsu, S., Albert, I., Rubinfeld, B. and **Polakis, P.** 1996 Deletion of amino-terminal structure stabilizes β-catenin in vivo and promotes the hyperphosphorylation of the APC tumor suppressor protein. **Mol. Cell. Biol.**16, 4088-4094.

44. Hart, M. J., Callow, M. G., Sousa, B. and **Polakis P.** 1996 IQGAP1, a calmodulin binding protein witha rasGAP related domain, is a potential effector for cdc42Hs. **EMBO J.** 15, 2997-3005.

45. Nathke, I. S., Adams, C. L., **Polakis, P.**, Sellin, J. and Nelson, W. J. 1996 The adenomatous polyposis coli (APC) tumor suppressor protein is localized to plasma membrane sites involved in active epithelial cell migration. **J. Cell. Biol.** 134, 165-180.

46. Hart, M. J., Sharma, S., elMasry, N., Qui, R-G., McCabe, P., **Polakis, P.** and Bollag, G. 1996 Identification of a novel guanine nucleotide exchange factor for the rho GTPase. **J. Biol. Chem.** 271, 25452.

47. Thomas JE, Smith M, Rubinfeld B, Gutowski M, Beckmann RP, and **Polakis P.** 1996 Subcellular localization and analysis of apparent 180-kDa and 220-kDa proteins of the breast cancer susceptibility gene, BRCA1. **J. Biol. Chem.** 1996 271, 28630-28635

48. Hayashi, S., Rubinfeld, B., Souza, B., **Polakis, P.**, Wieschaus, E., and Levine, A. 1997 A Drosophila homolog of the tumor suppressor adenomatous polyposis coli

down-regulates β -catenin b its zygotic expression is not esser for the regulation of armadillo. **Proc. Natl. Acad. Sci.** 94, 242-247.

49. Vleminckx, K., Rubinfeld, B., **Polakis, P.** and Gumbiner, B. 1997 The APC tumor suppressor protein induces a new axis in Xenopus embryos. **J. Cell. Biol.** 136, 411-420.

50. Rubinfeld, B., Robbins, P., El-Gamil, M., Albert, I., Porfiri, P. and **Polakis, P.** 1997 Stabilization of β-catenin by genetic defects in melanoma cell lines. **Science** 275, 1790-1792.

51. Polakis, P. The adenomatous polyposis coli (APC) tumor suppressor. 1997 **Biochem. Biophys. Acta**, 1332, F127-F147.

52. Rubinfeld, B., Albert, I., Porfiri, E., Munemitsu, S., and **Polakis, P** 1997 Loss of β-catenin regulation by the APC tumor suppressor protein correlates with loss of structure due to common somatic mutations of the gene. **Cancer Res.** 57, 4624-4630.

53. Porfiri, E., Rubinfeld, B., Albert, I., Hovanes. K., Waterman, M., and **Polakis, P.** 1997 Induction of a β-catenin-LEF-1 complex by wnt-1 and transforming mutants of β-catenin. **Oncogene** 15, 2833-2839.

54. Thomas JE, Smith M, Tonkinson JL, Rubinfeld B, and **Polakis P.**, 1997 Induction of phosphorylation on BRCA1 during the cell cycle and after DNA damage. **Cell Growth Differ.** 8, 801-809.

55. Hart, M., de los Santos, R., Albert, I., Rubinfeld, B., and **Polakis P.**, 1998 Down regulation of β-catenin by human Axin and its association with the adenomatous polyposis coli (APC) tumor suppressor, β-catenin and glycogen synthase kinase 3β. **Current Biology 8,** 573-581.

56. Polakis, P. 1998 The oncogenic activation of β-catenin. **Current Opinions in Genetics and Development** 9, 15-21

57. Matt Hart, Jean-Paul Concordet, Irina Lassot, Iris Albert, Rico del los Santos, Herve Durand, Christine Perret, Bonnee Rubinfled, Florence Margottin, Richard Benarous and **Paul Polakis**. 1999 The F-box protein β–TrCP associates with phosphorylated β–catenin and regulates its activity in the cell. **Current Biology** 9, 207-10.

58. Howard C. Crawford, Barbara M. Fingleton, Bonnee Rubinfeld, **Paul Polakis** and Lynn M. Matrisian 1999 The metalloproteinase matrilysin is a target of β–catenin transactivation in intestinal tumours. **Oncogene** 18, 2883-91.

59. Meng J, Glick JL, **Polakis P**, Casey PJ. 1999 Functional interaction between Galpha(z) and Rap1GAP suggests a novel form of cellular cross-talk. **J Biol Chem.** 17, 36663-9

60. Vijayasurian Easwaran, ginia Song, **Paul Polakis** and Ste Byers 1999 The ubiquitin-proteosome pathway and serine kinase activity modulate APC mediated regulation of β-catenin-LEF signaling. **J. Biol. Chem.** 274(23):16641-5.

61 Polakis P, Hart M and Rubinfeld B. 1999 Defects in the regulation of beta-catenin
in colorectal cancer. **Adv Exp Med Biol.** 470, 23-32

62 Shen Z, Batzer A, Koehler JA, **Polakis P**, Schlessinger J, Lydon NB, Moran MF. 1999 Evidence for SH3 domain directed binding and phosphorylation of Sam68 by Src. **Oncogene.** 18, 4647-53

64. Thomas GM, Frame S, Goedert M, Nathke I, **Polakis P**, Cohen P. 1999 A GSK3- binding peptide from FRAT1 selectively inhibits the GSK3-catalysed phosphorylation of axin and beta-catenin. **FEBS Lett.** 458, 247-51.

65. Peifer M, **Polakis P.** 2000 Wnt signaling in oncogenesis and embryogenesis--a look outside the nucleus. **Science** 287,1606-9.

66. **Polakis P.** 2000 Wnt signaling and cancer. **Genes Dev**;14, 1837-1851.

67. Spink KE, **Polakis P**, Weis WI 2000 Structural basis of the Axin-adenomatous polyposis coli interaction. **EMBO J** 19, 2270-2279.

68. Szeto , W., Jiang, W., Tice, D.A., Rubinfeld, B., Hollingshead, P.G., Fong, S.E., Dugger, D.L., Pham, T., Yansura, D.E., Wong, T.A., Grimaldi, J.C., Corpuz, R.T., Singh J.S., Frantz, G.D., Devaux, B., Crowley, C.W., Schwall, R.H., Eberhard, D.A.,
Rastelli, L., **Polakis, P.** and Pennica, D. 2001 Overexpression of the Retinoic Acid-
Responsive Gene Stra6 in Human Cancers and its Synergistic Induction by Wnt-1 and
Retinoic Acid. **Cancer Res** 61, 4197-4204.

69. Rubinfeld B, Tice DA, **Polakis P.** 2001 Axin dependent phosphorylation of the adenomatous polyposis coli protein mediated by casein kinase 1 epsilon. **J Biol Chem**
276, 39037-39045.

70. **Polakis P.** 2001 More than one way to skin a catenin. **Cell** 2001 105, 563-566.

71. Tice DA, Soloviev I, **Polakis P.** 2002 Activation of the Wnt Pathway Interferes withSerum Response Element-driven Transcription of Immediate Early Genes. **J Biol.**
Chem. 277, 6118-6123.

72. Tice DA, Szeto W, Soloviev I, Rubinfeld B, Fong SE, Dugger DL, Winer J,

Williams PM, Wieand D, Sr) V, Schwall RH, Pennnica D, **Pol: ; P.** 2002 Synergistic activation of tumor antigens by wnt-1 signaling and retinoic acid revealed by gene expression profiling. **J Biol Chem.** 277,14329-14335.

73. **Polakis, P.** 2002 Casein kinase I: A wnt'er of disconnect. **Curr. Biol.** 12, R499.

74. Mao,W. , Luis, E., Ross, S., Silva, J., Tan, C., Crowley, C., Chui, C., Franz, G., Senter, P., Koeppen, H., **Polakis, P.** 2004 EphB2 as a therapeutic antibody drug target for the treatment of colorectal cancer. **Cancer Res.** 64, 781-788.

75. Shibamoto, S., Winer, J., Williams, M., Polakis, P. 2003 A Blockade in Wnt signaling is activated following the differentiation of F9 teratocarcinoma cells. **Exp. Cell Res.** 29211-20.

76. Zhang Y, Eberhard DA, Frantz GD, Dowd P, Wu TD, Zhou Y, Watanabe C, Luoh SM, **Polakis P,** Hillan KJ, Wood WI, Zhang Z. 2004 GEPIS--quantitative gene expression profiling in normal and cancer tissues. **Bioinformatics**, April 8

EXHIBIT B

	tumor mRNA	tumor IHC
UNQ2525	+	+
UNQ2378	+	+
UNQ972	+	-
UNQ97671	+	+
UNQ2964	+	+
UNQ323	+	+
UNQ1655	+	+
UNQ2333	+	+
UNQ9638	+	+
UNQ8209	+	+
UNQ6507	+	+
UNQ8196	+	+
UNQ9109	+	+
UNQ100	+	+
UNQ178	+	+
UNQ1477	+	+
UNQ1839	+	+
UNQ2079	+	+
UNQ8782	+	+
UNQ9646	+	-
UNQ111	+	+
UNQ3079	+	+
UNQ8175	+	+
UNQ9509	+	+
UNQ10978	+	-
UNQ2103	+	+
UNQ1563	+	+
UNQ16188	+	+
UNQ13589	+	+
UNQ1078	+	+
UNQ879	+	+

1: Cancer Res. 1995 Mar 1;55(5):1168-75.

Malignant transformation of the human endometrium is associated with overexpression of lactoferrin messenger RNA and protein.

Walmer DK, Padin CJ, Wrona MA, Healy BE, Bentley RC, Tsao MS, Kohler MF, McLachlan JA, Gray KD.

Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, North Carolina 27710.

In the mouse uterus, lactoferrin is a major estrogen-inducible uterine secretory protein, and its expression correlates directly with the period of peak epithelial cell proliferation. In this study, we examine the expression of lactoferrin mRNA and protein in human endometrium, endometrial hyperplasias, and adenocarcinomas using immunohistochemistry, Western immunoblotting, and Northern and in situ RNA hybridization techniques. Our results reveal that lactoferrin is expressed in normal cycling endometrium by a restricted number of glandular epithelial cells located deep in the zona basalis. Two thirds (8 of 12) of the endometrial adenocarcinomas examined overexpress lactoferrin. This tumor-associated increase in lactoferrin expression includes an elevation in the mRNA and protein of individual cells and an increase in the number of cells expressing the protein. In comparison, only 1 of the 10 endometrial hyperplasia specimens examined demonstrates an increase in lactoferrin. We also observe distinct cytoplasmic and nuclear immunostaining patterns under different fixation conditions in both normal and malignant epithelial cells, similar to those previously reported in the mouse reproductive tract. Serial sections of malignant specimens show a good correlation between the localization of lactoferrin mRNA and protein in individual epithelial cells by in situ RNA hybridization and immunohistochemistry. Although the degree of lactoferrin expression in the adenocarcinomas did not correlate

with the tumor stage, grade, or depth of invasion in these 12 patients, there was a striking inverse correlation between the presence of progesterone receptors and lactoferrin in all 8 lactoferrin-positive adenocarcinomas. In summary, lactoferrin is expressed in a region of normal endometrium known as the zona basalis which is not shed with menstruation and is frequently overexpressed by progesterone receptor-negative cells in endometrial adenocarcinomas.

PMID: 7867003 [PubMed - indexed for MEDLINE]

About Entrez
NCBI Toolbar
Text Version
Entrez PubMed
Overview
Help | FAQ
Tutorials
New/Noteworthy
E-Utilities
PubMed Services
Journals Database
MeSH Database
Single Citation Matcher
Batch Citation Matcher
Clinical Queries
Special Queries
LinkOut
My NCBI
Related Resources
Order Documents
NLM Mobile
NLM Catalog
NLM Gateway
TOXNET
Consumer Health
Clinical Alerts
ClinicalTrials.gov
PubMed Central

1: Tumour Biol. 2004 Jul-Aug;25(4):161-71.

Related Articles, Links

KARGER Full Text

Alteration of frizzled expression in renal cell carcinoma.

Janssens N, Andries L, Janicot M, Perera T, Bakker A.

Department of Biochemistry, University of Antwerp, Wilrijk, Belgium. njansse9@prdbe.jnj.com

To evaluate the involvement of frizzled receptors (Fzds) in oncogenesis, we investigated mRNA expression levels of several human Fzds in more than 30 different human tumor samples and their corresponding (matched) normal tissue samples, using real-time quantitative PCR. We observed that the mRNA level of Fzd5 was markedly increased in 8 of 11 renal carcinoma samples whilst Fzd8 mRNA was increased in 7 of 11 renal carcinoma samples. Western blot analysis of crude membrane fractions revealed that Fzd5 protein expression in the matched tumor/normal kidney samples correlated with the observed mRNA level. Wnt/beta-catenin signaling pathway activation was confirmed by the increased expression of a set of target genes. Using a kidney tumor tissue array, Fzd5 protein expression was investigated in a broader panel of kidney tumor samples. Fzd5 membrane staining was detected in 30% of clear cell carcinomas, and there was a strong correlation with nuclear cyclin D1 staining in the samples. Our data suggested that altered expression of certain members of the Fzd family, and their downstream targets, could provide alternative mechanisms leading to activation of the Wnt signaling pathway in renal carcinogenesis. Fzd family members may have a role as a biomarker.

PMID: 15557753 [PubMed - indexed for MEDLINE]

May 22 2006 06:31:57

Limits: **Publication Date from 1992 to 1992**

About Entrez
NCBI Toolbar
Text Version

Entrez PubMed
Overview
Help | FAQ
Tutorials
New/Noteworthy
E-Utilities

PubMed Services
Journals Database
MeSH Database
Single Citation Matcher
Batch Citation Matcher
Clinical Queries
Special Queries
LinkOut
My NCBI

Related Resources
Order Documents
NLM Mobile
NLM Catalog
NLM Gateway
TOXNET
Consumer Health
Clinical Alerts
ClinicalTrials.gov
PubMed Central

1: Breast Cancer Res Treat. 1992;24(1):71-4.

Related Articles, Links

Expression of the pS2 gene in breast tissues assessed by pS2-mRNA analysis and pS2-protein radioimmunoassay.

Hahnel E, **Robbins P**, **Harvey J**, **Sterrett G**, **Hahnel R**.

Department of Pathology, University of Western Australia, Queen Elizabeth II Medical Centre, Nedlands.

The expression of the pS2 gene in breast tissues was assessed by measuring pS2-protein using a radioimmunoassay, and by determining pS2-mRNA using Northern blotting. There was a good correlation between the two measurements, indicating that expression of the pS2 gene in breast tissues may be assessed by either method. Since radioimmunoassay is technically easier and more efficient than Northern blotting, radioimmunoassay will be the method of choice in routine applications.

PMID: 1463873 [PubMed - indexed for MEDLINE]

About Entrez
NCBI Toolbar
Text Version

Entrez PubMed
Overview
Help | FAQ
Tutorials
New/Noteworthy
E-Utilities

PubMed Services
Journals Database
MeSH Database
Single Citation Matcher
Batch Citation Matcher
Clinical Queries
Special Queries
LinkOut
My NCBI

Related Resources
Order Documents
NLM Mobile
NLM Catalog
NLM Gateway
TOXNET
Consumer Health
Clinical Alerts
ClinicalTrials.gov
PubMed Central

1: Int J Oncol. 2005 Nov;27(5):1257-63.

Related Articles, Links

Expression of human telomerase reverse transcriptase gene and protein, and of estrogen and progesterone receptors, in breast tumors: preliminary data from neo-adjuvant chemotherapy.

Kammori M, Izumiyama N, Hashimoto M, Nakamura K, Okano T, Kurabayashi R, Naoki H, Honma N, Ogawa T, Kaminishi M, Takubo K.

Division of Breast and Endocrine Surgery, Department of Surgery, Graduate School of Medicine, The University of Tokyo, Japan. kanmori-dis@umin.ac.jp

Human telomerase reverse transcriptase (hTERT), the catalytic subunit of telomerase, is very closely associated with telomerase activity. Telomerase has been implicated in cellular immortalization and carcinogenesis. In situ detection of hTERT will aid in determining the localization of telomerase-positive cells. The aim of this study was to detect expression of hTERT mRNA, hTERT protein, estrogen receptor (ER) and progesterone receptor (PR) in paraffin-embedded breast tissue samples and to investigate the relationship between hTERT expression and various clinicopathological parameters in breast tumorigenesis. We used in situ hybridization (ISH) to examine hTERT gene expression, and immunohistochemistry (IHC) to examine expression of hTERT protein, ER and PR, in breast tissues including 64 adenocarcinomas, 2 phyllode tumors and their adjacent normal breast tissues. hTERT gene expression was detected by ISH in 56 (88%) carcinomas, but in neither of the 2 phyllode tumors. hTERT protein expression was detected by IHC in 52 (81%) carcinomas, but in neither of the 2 phyllode tumors. Moreover, ER and PR were expressed in 42 (66%) and 42 (66%) carcinomas, respectively, and in neither of the 2 phyllode tumors. In 4 cases of breast carcinoma that strongly expressed hTERT gene and protein before treatment, neo-adjuvant chemotherapy led to disappearance of gene and protein expression in all cases. There was

a strong correlation between detection of hTERT gene expression by ISH and of hTERT protein by ICH in tissue specimens from breast tumors. These results suggest that detection of hTERT protein by ICH can be used to distinguish breast cancers as a potential diagnostic and therapeutic marker.

PMID: 16211220 [PubMed - indexed for MEDLINE]

Display Abstract Show 20 Sort by Send to

May 22 2006 06:31:57

[CANCER RESEARCH 61, 2409–2412, March 15, 2001]

Advances in Brief

BMI-1 Gene Amplification and Overexpression in Hematological Malignancies Occur Mainly in Mantle Cell Lymphomas[1]

Sílvia Beà, Frederic Tort, Magda Pinyol, Xavier Puig, Luis Hernández, Silvia Hernández, Pedro L. Fernández, Maarten van Lohuizen, Dolors Colomer, and Elias Campo[2]

The Hematopathology Section, Laboratory of Anatomic Pathology, Hospital Clínic, Institut d'Investigacions Biomèdiques "August Pi i Sunyer" (IDIBAPS), University of Barcelona, 08036 Barcelona, Spain [S. B., F. T., M. P., X. P., L. H., S. H., P. L. F., D. C., E. C.], and Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, 1066 CX Amsterdam, Netherlands [M. v. L.]

Abstract

The *BMI-1* gene is a putative oncogene belonging to the Polycomb group family that cooperates with c-*myc* in the generation of mouse lymphomas and seems to participate in cell cycle regulation and senescence by acting as a transcriptional repressor of the *INK4a/ARF* locus. The *BMI-1* gene has been located on chromosome 10p13, a region involved in chromosomal translocations in infant leukemias, and amplified in occasional non-Hodgkin's lymphomas (NHLs) and solid tumors. To determine the possible alterations of this gene in human malignancies, we have examined 160 lymphoproliferative disorders, 13 myeloid leukemias, and 89 carcinomas by Southern blot analysis and detected *BMI-1* gene amplification (3- to 7-fold) in 4 of 36 (11%) mantle cell lymphomas (MCLs) with no alterations in the *INK4a/ARF* locus. BMI-1 and $p16^{INK4a}$ mRNA and protein expression were also studied by real-time quantitative reverse transcription-PCR and Western blot, respectively, in a subset of NHLs. BMI-1 expression was significantly higher in chronic lymphocytic leukemia and MCL than in follicular lymphoma and large B cell lymphoma. The four tumors with gene amplification showed significantly higher mRNA levels than other MCLs and NHLs with the *BMI-1* gene in germline configuration. Five additional MCLs also showed very high mRNA levels without gene amplification. A good correlation between BMI-1 mRNA levels and protein expression was observed in all types of lymphomas. No relationship was detected between BMI-1 and $p16^{INK4a}$ mRNA levels. These findings suggest that *BMI-1* gene alterations in human neoplasms are uncommon, but they may contribute to the pathogenesis in a subset of malignant lymphomas, particularly of mantle cell type.

Introduction

The *BMI-1*[3] gene is a putative oncogene of the Polycomb group originally identified by retroviral insertional mutagenesis in Eμ-c-myc transgenic mice infected with the Moloney murine leukemia virus (1, 2). These animals had a rapid development of pre-B cell lymphomas showing frequent proviral insertions near the *BMI-1* gene. This integration resulted in BMI-1 overexpression suggesting a cooperative effect between *C-MYC* and *BMI-1* genes in the development of these tumors (3, 4). Recent studies have indicated that the *BMI-1* gene may also participate in cell cycle control and senescence through the *INK4a/ARF* locus by acting as an upstream negative regulator of $p16^{INK4a}$ and $p14/p19^{ARF}$ gene expression (5). The human *BMI-1* gene has been mapped to chromosome 10p13 (6), a region involved in chromosomal translocations in infant leukemias (7) and rearrangements in malignant T cell lymphomas (8, 9). More recently, high-level DNA amplifications of this region have been found by comparative genomic hybridization in NHLs and solid tumors (10, 11). However, the possible implication of the *BMI-1* gene in these alterations and its role in the pathogenesis of human tumors is not known. The aim of this study was to analyze the possible *BMI-1* gene alterations and expression in a large series of human neoplasms and to determine the relationship with *INK4a/ARF* locus aberrations.

Materials and Methods

Case Selection. A series of 262 human tumors, including 173 hematological malignancies and 89 carcinomas (Table 1), matched normal tissues from all carcinomas, 11 samples of normal peripheral mononuclear cells, and 5 reactive lymph nodes and tonsils, were selected based on the availability of frozen samples for molecular analysis.

DNA Extraction and Southern Blot Analysis. Genomic DNA was obtained using Proteinase K/RNase treatment. 15 μg were digested with *Eco*RI and *Hin*dIII restriction enzymes (Life Technologies, Inc., Gaithersburg, MD), for Southern blot analysis and hybridized with a 1.5-kb *Pst*I fragment of the partial *BMI-1* cDNA (6).

RNA Extraction and Real-time Quantitative RT-PCR. Total RNA was obtained from 67 lymphoid neoplasms (10 CLLs, 27 MCLs, 8 FLs, and 22 LCLs) using guanidine/isothiocyanate extraction and cesium/chloride gradient centrifugation. One μg of total RNA was transcribed into cDNA using MMLV-reverse transcriptase (Life Technologies, Inc.) and random hexamers, following manufacturer's directions. Sequences of the *BMI-1* and the *p16* detection probes and primers were designed using the Primer Express program (Applied Biosystems, Foster City) as follows: *BMI-1* sense, 5′-CTGGTTGCCCATTGACAGC-3′; *BMI-1* antisense, 5′-CAGAAAATGAATGCGAGCCA-3′; *p16* sense, 5′-CAACGCACCGAATAGTTACGG-3′; *p16* antisense, 5′-AACTTCGTCCTCCAGAGTCGC-3′. The probes *BMI-1*, 5′-CAGCTCGCTTCAAGATGGCCGC-3′, and *p16*, 5′-CGGAGGCCGATCCAGGTGGGTA-3′, were labeled with 6-carboxy-fluorescein as the reporter dye. The TaqMan-GAPDH Control Reagents (Applied Biosystems) were used to amplify and detect the *GAPDH* gene, as recommended by the manufacturer. The quantitative assay amplified 1 μml of cDNA in two to four replicates using the primers and probes described above and the standard master mix (Applied Biosystems). All reactions were performed in an ABI PRISM 7700 Sequence Detector System (Applied Biosystems). GAPDH, BMI-1, and $p16^{INK4a}$ expression was related to a standard curve derived from serial dilutions of Raji cDNA. The RUs of BMI-1 and $p16^{INK4a}$ expression were defined as the mRNA levels of these genes normalized to the GADPH expression level in each case.

Protein Analysis. Whole-cell protein extracts were obtained from additional frozen tissue available in 31 cases (7 CLLs, 12 MCLs, 8 FLs, and 4 LCLs), loaded onto a 10% SDS-polyacrylamide gel, and electroblotted to a nitrocellulose membrane (Amersham). Blocked membranes were incubated sequentially with the monoclonal antibody BMI-F6 (12), antimouse conju-

Received 10/16/00; accepted 1/29/01.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported by Grant SAF 99/20 from Comision Interministerial de Ciencia y Tecnologia, European Union Contract QLG1-CT-2000-689, the Asociación Española contra el Cáncer, and Generalitat de Catalunya 98SGR21. S. B. and F. T. were fellows supported by Spanish Ministerio de Educación y Cultura, and S. H. was supported by the Asociación Española contra el Cáncer and the Fundació Rius i Virgili.

[2] To whom requests for reprints should be addressed, at the Department of Pathology, Hospital Clínic, University of Barcelona, Villarroel 170, 08036-Barcelona, Spain. Phone: 34 93 227 5450; Fax: 34 93 227 5572; E-mail: campo@medicina.ub.es.

[3] The abbreviations used are: BMI-1, B cell-specific Moloney murine leukemia virus integration site 1; NHL, non-Hodgkin's lymphoma; CLL, chronic lymphocytic leukemia; FL, follicular lymphoma; LCL, large B cell lymphoma; MCL, mantle cell lymphoma; RT-PCR, reverse-transcription-PCR; RU, relative units.

Table 1 *Hematological malignancies and solid tu... amples analyzed for BMI-1 gene alterations*

Tissue samples	No. of cases
Hematological malignancies	
Hodgkin's disease	2
B cell lymphoproliferative disorders	
B-Acute lymphoblastic leukemia	14
CLL	29
Hairy cell leukemia	4
FL	15
MCL	36
LCL	40
T cell lymphoproliferative disorders	
T-Acute lymphoblastic leukemia	8
Large granular cell leukemia	4
Peripheral T-cell lymphoma	8
Myeloproliferative disorders	
Acute myeloid leukemia	7
Chronic myeloid leukemia	6
Solid tumors	
Colon carcinoma	26
Breast carcinoma	29
Laryngeal squamous cell carcinoma	34
Total	262

gated to horseradish peroxidase (Amersham), and detected by enhanced chemiluminescence (Amersham) according to the manufacturer's recommendations.

Statistical Analysis. Because of the non-normal distribution of the samples and the small size of some subsets of tumors, the statistical evaluation was performed using nonparametric tests (SPSS, version 9.0). Comparison between mRNA expression levels in the different groups of NHLs was performed using the Kruskal-Wallis Test, with a P for significance set at 0.05. For differences between particular groups, the conservative Bonferroni procedure was performed, and the P was set at 0.005. The remaining statistical analyses were carried out using the Mann-Whitney nonparametric U test (significance, $P < 0.05$). The comparison between BMI-1 and $p16^{INK4a}$ quantitative mRNA levels was also performed using the Pearson's correlation coefficient.

Results

***BMI-1* Gene Amplification.** The *BMI-1* gene was examined by Southern blot in a large series of human tumors and normal samples (Table 1). The cDNA probe used in the study detected three *Eco*RI fragments of 7.3, 3.8, and 2.6 kb and three *Hin*dIII fragments of 6.2, 4, and 3.5 kb. *BMI-1* gene amplification (3- to 7-fold) was detected in 4 of 36 (11%) MCLs (Fig. 1). The amplifications were confirmed with both restriction enzymes. The amplified MCLs were two blastoid and two typical variants. No amplifications were observed in any of the solid tumors when compared with their respective matched non-neoplastic mucosa. No *BMI-1* gene rearrangements were observed in any of the samples examined.

BMI-1 mRNA Expression. To determine the BMI-1 expression pattern in NHL we analyzed BMI-1 mRNA levels by real-time quantitative RT-PCR in 67 lymphomas (10 CLLs, 27 MCLs, 8 FLs, and 22 LCLs), including the four tumors with gene amplification. A distinct BMI-1 mRNA expression pattern was observed in the different types of lymphomas (Fig. 2; Kruskal-Wallis Test; $P < 0.001$). The BMI mRNA levels in CLLs (mean, 2.2 RU; SD, 1.3) and MCLs with no *BMI-1* gene amplification (mean, 2.5 RU; SD, 2.3) were significantly higher than in FLs (mean, 0.9 RU; SD, 0.8) and LCLs (mean, 0.6 RU; SD, 0.4; Mann-Whitney nonparametric U test; $P < 0.01$). The 4 MCLs with *BMI-1* gene amplification showed significantly higher levels of expression than all other groups of tumors (mean, 5.1 RU; SD, 1.6; $P < 0.005$). In addition, five typical MCLs with no structural alterations of the gene also showed very high levels of BMI-1 mRNA expression ranging from 4 to 9.8 RU, similar to cases with gene amplification (Fig. 2*A*).

BMI-1 Protein Expression. BMI-1 protein expression was examined by Western blot in 31 tumors (7 CLLs; 12 MCLs, including two cases with *BMI-1* gene ...plification and 4 cases with mRNA overexpression and no structural alteration of the gene; 8 FLs, and 4 LCLs) in which additional frozen tissue was available. The monoclonal antibody against BMI-1 detected three closely migrating proteins of M_r 45,000–48,000 (2). The two more slowly migrating bands probably represent phosphorylated isoforms of the protein (12). The two MCLs with gene amplification and three of four cases with mRNA overexpression without amplification of the gene showed very high levels of protein expression. The remaining MCLs and CLLs showed intermediate levels of protein expression, whereas low- or no-expression signals were detected in the LCLs and FLs included in the study (Fig. 3). These results indicate that BMI-1 protein expression in NHL is concordant with the mRNA levels observed by real-time quantitative RT-PCR.

Relationship between BMI-1 and $p16^{INK4a}$ Gene Alterations. The *INK4a/ARF* locus has been recently identified as a downstream target of the transcriptional repressing activity of the *BMI-1* gene, suggesting that this gene may contribute to human neoplasias with wild type *INK4/ARF* (5). Most of the lymphoproliferative disorders analyzed in the present study, including the four cases with *BMI-1* gene amplification, had been previously examined for *p53* gene mutations and *INK4a/ARF* locus alterations, including gene deletions, mutations, hypermethylation, and expression (13, 14). The four MCLs with *BMI-1* gene amplification and mRNA overexpression and the five tumors with BMI-1 mRNA overexpression with no structural alterations of the gene showed a wild-type configuration of the *INK4a/ARF* locus (13). However, one case with *BMI-1* gene amplification and one case with mRNA overexpression with no alteration of the gene showed *p53* gene mutations associated with allelic deletions.

To determine the possible relationship between BMI-1 and $p16^{INK4a}$ mRNA expression, $p16^{INK4a}$ mRNA levels were evaluated by real-time quantitative RT-PCR in 50 tumors (10 CLLs, 27 MCLs, and 13 LCLs), including 6 cases with alterations in the *INK4a/ARF* locus (2 MCLs and 1 LCL with $p16^{INK4a}$ gene deletion, 2 LCLs with *p16* promoter hypermethylation, and 1 CLL with $p16^{INK4a}$ gene mutation), and the 4 lymphomas with *BMI-1* amplification. Negative or negligible levels of $p16^{INK4a}$ were observed in the 6 tumors with *INK4a/ARF* locus alterations. These cases were not included in the comparisons between BMI-1 and $p16^{INK4a}$ mRNA expression. The $p16^{INK4a}$ expression levels were relatively similar in the different types of tumors. Only LCLs tended to have lower levels of expression, but the differences did not reach statistical significance (Fig. 2*B*). No differences were observed in the $p16^{INK4a}$ mRNA levels between tumors with *BMI-1* gene amplification and overexpression and lymphomas with germline configuration of the gene.

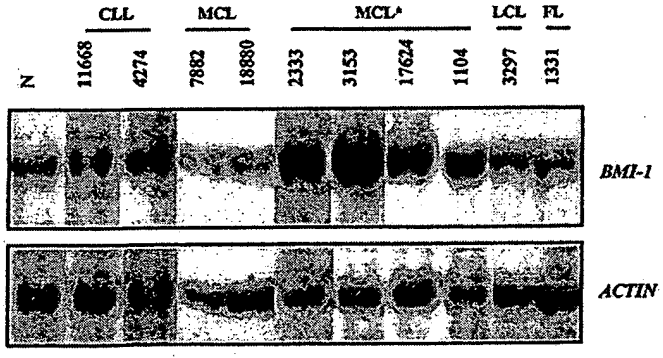

Fig. 1. Southern blot analysis of *BMI-1* gene. Four MCLs (*MCL**) showed *BMI-1* gene amplification (3- to 7-fold) compared with non-neoplastic tissues (*N*) and other NHLs. No amplifications or gene rearrangements were detected in the remaining NHLs and carcinomas included in the study.

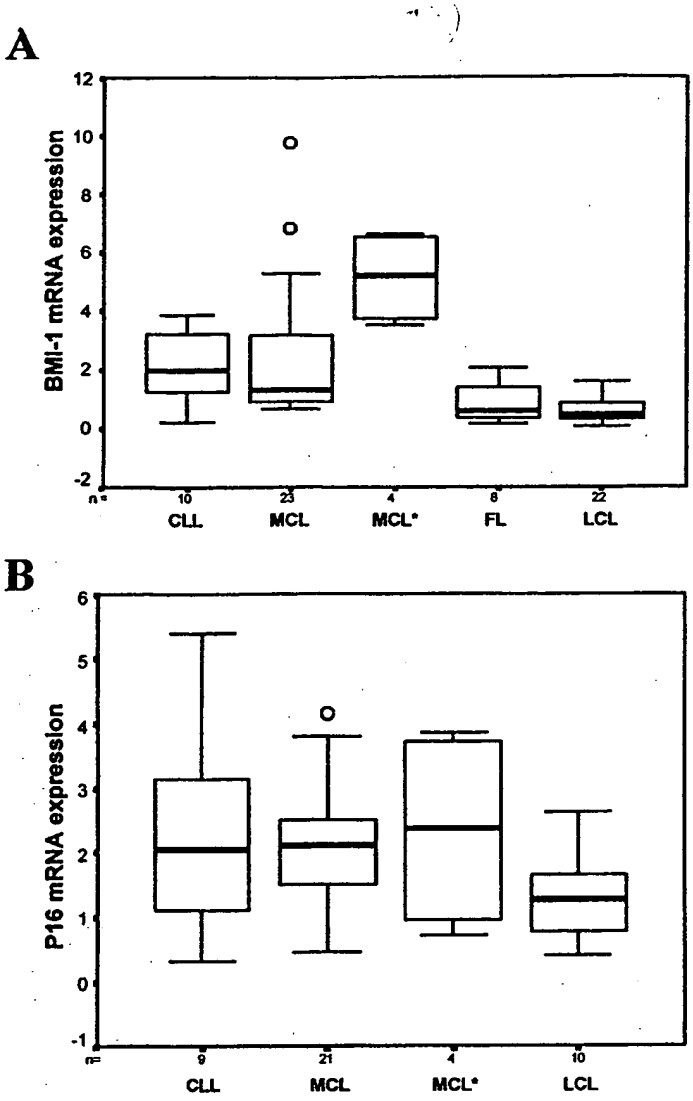

Fig. 2. *A*, quantitative BMI-1 mRNA transcript analysis (median and range) using real-time RT-PCR in a series of NHLs. MCLs with *BMI-1* gene amplification (*MCL**) revealed significantly higher overall BMI-1 mRNA levels than all other types of NHLs, including MCLs with no structural alterations of the gene ($P < 0.005$). MCLs and CLLs expressed significantly higher levels than FLs and LCLs ($P < 0.001$). Results are depicted as the ratio of absolute BMI-1:GADPH mRNA transcript numbers (RU). *Bars*, SD. *B*, quantitative $p16^{INK4a}$ mRNA transcript analysis (median and range) using real-time RT-PCR in a series of NHLs. Expression levels were relatively similar in the different types of tumors. Results are depicted as the ratio of absolute $p16^{INK4a}$:GADPH mRNA transcript numbers (RU). *Bars*, SD.

Discussion

In the present study, we have examined a large series of human tumors for the presence of gene alterations and mRNA expression of the *BMI-1* gene. Gene amplification was identified in four MCLs. These tumors showed significantly higher levels of mRNA and protein expression compared with other lymphomas with *BMI-1* in germline configuration. BMI-1 expression levels were also highly up-regulated in a subset of MCLs with no apparent structural alterations of the gene. No alterations were detected in any of the different types of carcinomas included in the study. *BMI-1* is considered an oncogene belonging to the Polycomb group family of genes. These proteins mainly act as transcriptional regulators, controlling specific target genes involved in development, cell differentiation, proliferation, and senescence. Different studies have shown the implication of BMI-1 overexpression in the development of lymphomas in murine and feline animal models (3, 4). The findings of the present study indicate for the first time that *B-* gene alterations in human neoplasms are an uncommon phenomenon, but they seem to occur mainly in a subset of NHLs, particularly of mantle cell type.

The human *BMI-1* gene has been mapped to chromosome 10p13. High-level DNA amplifications and gains in this region have been identified by comparative genomic hybridization in occasional solid tumors and NHLs (10, 11). Different chromosomal translocations involving the 10p13 region have also been identified in infant leukemias and T cell lymphoproliferative disorders (7, 8, 15). Most acute leukemias with this chromosomal alteration occur in children <12 months of age, whereas it seems to be extremely rare in adults. 10p translocations in T-cell lymphoproliferative disorders have been observed mainly in adult T cell leukemia/lymphomas and occasional cutaneous T cell lymphomas. In our study, we did not observe *BMI-1* rearrangements or amplifications in any of the acute leukemias or T cell lymphomas. However, all of the acute leukemias in this study were diagnosed in patients over 16 years, and no adult T cell leukemia/lymphomas or cutaneous lymphomas could be included in the series. Similarly, high-level DNA amplifications at the 10p13 region have been detected in head and neck carcinomas and other solid tumors. Although we found no evidence for *BMI-1* gene rearrangements or amplifications in a substantial set of carcinomas, this does not exclude the possibility of increased gene expression or protein levels in these tumors. Additional studies are required to elucidate the possible involvement of *BMI-1* in these particular groups of human neoplasms.

In human hematopoietic cells, BMI-1 is preferentially expressed in primitive CD34+ bone marrow cells, whereas it is negative or very low in more mature CD34− cells (16). In peripheral lymphocytes, and particularly in follicular B cells, BMI-1 protein expression has been detected in resting cells of the mantle zone, whereas it is down-regulated in proliferating germinal center cells (17, 18). These observations indicate that BMI-1 expression in normal hematopoietic cells is tightly regulated in relation with cell differentiation in bone marrow and antigen-specific response in peripheral lymphocytes. BMI-1 expression in human tumors has not been examined previously. In this study, we have demonstrated that BMI-1 mRNA and protein expression show a distinct pattern in different types of lymphomas. Thus, BMI-1 levels were low in LCLs and FLs and significantly higher in MCLs and CLLs. These findings suggest that BMI-1 expression patterns in B cell lymphomas maintain in part the expression profile of their normal cell counterparts; because FLs and at least a subgroup of LCLs are considered lymphomas derived from follicular germinal center cells, whereas MCLs and CLLs are tumors mainly derived from naive pregerminal center cells. However, the four MCLs with *BMI-1* gene amplification expressed significantly higher mRNA levels than all other tumors. In addition, five MCLs with no structural alterations of the gene showed high mRNA levels similar to those observed in tumors with *BMI-1* gene amplification, suggesting that other mechanisms may be involved in up-regulation of the gene in these lymphomas. Different studies using animal models have shown a dose-dependent effect of *BMI-1* gene expression on skeleton development

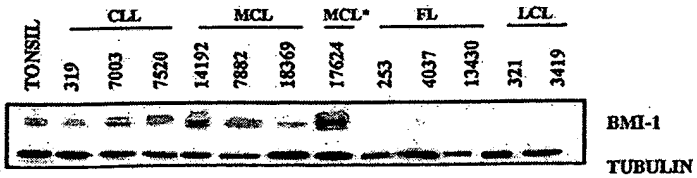

Fig. 3. Western blot analysis of BMI-1 protein in NHLs. The amplified MCL (17624) showed the highest BMI-1 protein levels, whereas other MCLs and CLLs had intermediate levels of expression. Very low or negative signal was observed in FLs and LCLs.

and lymphomagenesis (1, 3). These observations suggest that the high mRNA and protein levels detected in a subset of MCLs may play a role in the pathogenesis of these neoplasms.

Recent studies have identified the *INK4/ARF* locus as a downstream target of the BMI-1 transcriptional repressor activity, suggesting that BMI-1 overexpression may contribute to human neoplasias that retain the wild-type *INK4a/ARF* locus (5). Interestingly, in our study, *BMI-1* amplification and overexpression appeared in tumors with no alterations in $p16^{INK4a}$ and $p14^{ARF}$ genes. However, we could not detect differences in the expression levels of $p16^{INK4a}$ in tumors with and without *BMI-1* gene alterations. The reasons for this apparent discrepancy with experimental observations are not clear. One possibility may be that genes other than *INK4a/ARF* are the main targets of BMI-1 repressor activity in these tumors. Particularly, different genes of the HOX family are regulated by BMI-1 and may also be involved in lymphomagenesis (19, 20).

In conclusion, the findings of this study indicate that *BMI-1* gene expression is differentially regulated in B cell lymphomas. Alterations of the gene seem to be an uncommon phenomenon in human neoplasms, but they may contribute to the pathogenesis in a subset of MCLs. Although, *BMI-1* gene alterations occurred in tumors with wild-type *INK4a/ARF* locus, the possible cooperation between these genes and the oncogenic mechanisms of BMI-1 in human neoplasms require additional analysis.

Acknowledgments

The authors thank Iracema Nayach for her excellent technical assistance.

References

1. Haupt, Y., Alexander, W. S., Barri, G., Klinken, S. P., and Adams, J. M. Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in E μ-myc transgenic mice. Cell, *65:* 753–763, 1991.
2. van Lohuizen, M., Verbeek, S., Scheijen, B., Wientjens, E., van der Gulden, H., and Berns, A. Identification of cooperating oncogenes in E μ-myc transgenic mice by provirus tagging. Cell, *65:* 737–752, 1991.
3. Alkema, M. J., Jacobs, H., van Lohuizen, M., and Berns, A. Perturbation of B and T cell development and predisposition to lymphomagenesis in EμBmi1 transgenic mice require the Bmi1 RING finger. Oncogene, *15:* 899–910, 1997.
4. Haupt, Y., Bath, M. L., Harris, A. W., and Adams, J. M. *Bmi-1* transgene induces lymphomas and collaborates with myc in tumorigenesis. Oncogene, *8:* 3161–3164, 1993.
5. Jacobs, J. J., Kieboom, I. ..rino, S., DePinho, R. A., and van Lohuizen, M. The oncogene and Polycomb-group gene *bmi-1* regulates cell proliferation and senescence through the *ink4a* locus. Nature (Lond.), *397:* 164–168, 1999.
6. Alkema, M. J., Wiegant, J., Raap, A. K., Berns, A., and van Lohuizen, M. Characterization and chromosomal localization of the human proto-oncogene *BMI-1*. Hum. Mol. Genet., *2:* 1597–1603, 1993.
7. Pui, C. H., Raimondi, S. C., Murphy, S. B., Ribeiro, R. C., Kalwinsky, D. K., Dahl, G. V., Crist, W. M., and Williams, D. L. An analysis of leukemic cell chromosomal features in infants. Blood, *69:* 1289–1293, 1987.
8. Berger, R., Baranger, L., Bernheim, A., Valensi, F., Flandrin, G., and Berheimm, A. T. Cytogenetics of T-cell malignant lymphoma. Report of 17 cases and review of the chromosomal breakpoints. Cancer Genet. Cytogenet., *36:* 123–130, 1988.
9. D'Alessandro, E., Paterlini, P., Lo Re, M. L., Di Cola, M., Ligas, C., Quaglino, D., and Del Porto, G. Cytogenetic follow-up in a case of Sezary syndrome. Cancer Genet. Cytogenet., *45:* 231–236, 1990.
10. Bea, S., Ribas, M., Hernandez, J. M., Bosch, F., Pinyol, M., Hernandez, L., Garcia, J. L., Flores, T., Gonzalez, M., Lopez-Guillermo, A., Piris, M. A., Cardesa, A., Montserrat, E., Miro, R., and Campo, E. Increased number of chromosomal imbalances and high-level DNA amplifications in mantle cell lymphoma are associated with blastoid variants. Blood, *93:* 4365–4374, 1999.
11. Knuutila, S., Bjorkqvist, A. M., Autio, K., Tarkkanen, M., Wolf, M., Monni, O., Szymanska, J., Larramendy, M. L., Tapper, J., Pere, H., el-Rifai, W., Hemmer, S., Wasenius, V. M., Vidgren, V., and Zhu, Y. DNA copy number amplifications in human neoplasms: review of comparative genomic hybridization studies. Am. J. Pathol., *152:* 1107–1123, 1998.
12. Alkema, M. J., Bronk, M., Verhoeven, E., Otte, A., van't Veer, L. J., Berns, A., and van Lohuizen, M. Identification of Bmi1-interacting proteins as constituents of a multimeric mammalian polycomb complex. Genes Dev., *11:* 226–240, 1997.
13. Pinyol, M., Hernández, L., Martínez, A., Cobo, F., Hernández, S., Beà, S., López-Guillermo, A., Nayach, I., Palacín, A., Nadal, A., Fernández, P., Montserrat, E., Cardesa, A., and Campo, E. *INK4a/ARF* locus alterations in human non-Hodgkin's lymphomas mainly occur in tumors with wild type *p53* gene. Am. J. Pathol., *156:* 1987–1996, 2000.
14. Pinyol, M., Cobo, F., Beà, S., Jares, P., Nayach, I., Fernández, P. L., Montserrat, E., Cardesa, A., and Campo, E. *p16INK4a* gene inactivation by deletions, mutations, and hypermethylation is associated with transformed and aggressive variants of non-Hodgkin's lymphomas. Blood, *91:* 2977–2984, 1998.
15. Foot, A. B., Oakhill, A., and Kitchen, C. Acute monoblastic leukemia of infancy in Klinefelter's syndrome. Cancer Genet. Cytogenet., *61:* 99–100, 1992.
16. Lessard, J., Baban, S., and Sauvageau, G. Stage-specific expression of polycomb group genes in human bone marrow cells. Blood, *91:* 1216–1224, 1998.
17. Raaphorst, F. M., van Kemenade, F. J., Fieret, E., Hamer, K. M., Satijn, D. P., Otte, A. P., and Meijer, C. J. Cutting edge: polycomb gene expression patterns reflect distinct B cell differentiation stages in human germinal centers. J. Immunol., *164:* 1–4, 2000.
18. Raaphorst, F. M., van Kemenade, F. J., Blokzijl, T., Fieret, E., Hamer, K. M., Satijn, D. P., Otte, A. P., and Meijer, C. J. Coexpression of *BMI-1* and *EZH2* polycomb group genes in Reed-Sternberg cells of Hodgkin's disease. Am. J. Pathol., *157:* 709–715, 2000.
19. Gould, A. Functions of mammalian Polycomb group and trithorax group related genes. Curr. Opin. Genet. Dev., *7:* 488–494, 1997.
20. van Oostveen, J., Bijl, J., Raaphorst, F., Walboomers, J., and Meijer, C. The role of homeobox genes in normal hematopoiesis and hematological malignancies. Leukemia, *13:* 1675–1690, 1999.

American Journal of Pathology, Vol. 155, No. 3, September 1999

Id-1 and Id-2 Are Overexpressed in Pancreatic Cancer and in Dysplastic Lesions in Chronic Pancreatitis

Haruhisa Maruyama,* Jörg Kleeff,* Stefan Wildi,* Helmut Friess,† Markus W. Büchler,† Mark A. Israel,‡ and Murray Korc*

From the Division of Endocrinology, Diabetes, and Metabolism, Departments of Medicine, Biological Chemistry and Pharmacology, University of California, Irvine, California; the Department of Visceral and Transplantation Surgery,† University of Bern, Bern, Switzerland; and the Preuss Laboratory,‡ Department of Neurological Surgery, University of California, San Francisco, California*

Id proteins antagonize basic helix-loop-helix proteins, inhibit differentiation, and enhance cell proliferation. In this study we compared the expression of Id-1, Id-2, and Id-3 in the normal pancreas, in pancreatic cancer, and in chronic pancreatitis (CP). Northern blot analysis demonstrated that all three Id mRNA species were expressed at high levels in pancreatic cancer samples by comparison with normal or CP samples. Pancreatic cancer cell lines frequently coexpressed all three Ids, exhibiting a good correlation between Id mRNA and protein levels, as determined by immunoblotting with highly specific anti-Id antibodies. Immunohistochemistry using these antibodies demonstrated the presence of faint Id-1 and Id-2 immunostaining in pancreatic ductal cells in the normal pancreas, whereas Id-3 immunoreactivity ranged from weak to strong. In the cancer tissues, many of the cancer cells exhibited abundant Id-1, Id-2, and Id-3 immunoreactivity. Scoring on the basis of percentage of positive cells and intensity of immunostaining indicated that Id-1 and Id-2 were increased significantly in the cancer cells by comparison with the respective controls. Mild to moderate Id immunoreactivity was also seen in the ductal cells in the CP-like areas adjacent to these cells and in the ductal cells of small and interlobular ducts in CP. In contrast, in dysplastic and atypical papillary ducts in CP, Id-1 and Id-2 immunoreactivity was as significantly elevated as in the cancer cells. These findings suggest that increased Id expression may be associated with enhanced proliferative potential of pancreatic cancer cells and of proliferating or dysplastic ductal cells in CP. *(Am J Pathol 1999, 155:815–822)*

Basic helix-loop-helix (bHLH) proteins play an important role as transcription factors in cellular development, proliferation, and differentiation.[1,2] The basic domain of the bHLHs is required for binding to an E-box DNA sequence, thus promoting transcription of specific target genes. The HLH domain promotes dimer formation with various members of the bHLH protein family.[1,2] Homodimers of the class B family of bHLH proteins, including *Myo*D, *Neuro*D, and numerous other proteins, are known to activate tissue-specific genes.[3–5] These tissue-specific bHLHs typically form heterodimers with widely expressed class A bHLHs, which include proteins encoded by E2A, E2–2, HEB, and other genes (also termed E-proteins).[6–9] These heterodimers activate transcription of genes that are associated with differentiation.

Id genes encode a family of four HLH proteins that lack the basic DNA binding domain.[1,10] They act as dominant-negative HLH proteins by forming high affinity heterodimers with other bHLH proteins, thereby preventing them from binding to DNA and inhibiting transcription of differentiation-associated genes.[10–12] Id gene expression is down-regulated on differentiation in many cell types *in vitro* and *in vivo*.[13–18] In addition, Id proteins seem to be required for cell cycle progression through G_1/S phase in certain cell types, and interaction between Id-2 and pRB is associated with enhanced proliferation in some cell lines *in vitro*.[19–23]

Pancreatic cancer is the fifth leading cause of cancer death in the United States, with a mortality rate that virtually equals its incidence rate.[24] This malignancy is often associated with the overexpression of a variety of mitogenic growth factors and their receptors, and by oncogenic mutations of K-*ras* and inactivation of the p53 tumor suppressor gene.[25] We have recently reported that pancreatic cancers overexpress the HLH protein Id-2, and that enhanced expression of this protein is evident in the cytoplasm of the cancer cells within the pancreatic tumor mass.[26] It is not known, however, whether the expression of other Id proteins is altered in this malignancy, or whether their expression is altered in chronic pancreatitis

Contract grant sponsor: National Cancer Institute. Contract grant number: U. S. Public Health Service grant CA-40162.

Accepted for publication May 24, 1999.

Address reprint requests to Dr. Murray Korc, Division of Endocrinology, Diabetes and Metabolism, Medical Sciences I, C240, University of California, Irvine, CA 92697. E-mail: mkorc@uci.edu.

(CP), an inflammatory disease that is characterized by dysplastic ducts, foci of proliferating ductal cells, acinar cell degeneration, and fibrosis.[27] We now report that there is a five- to sixfold increase in Id-1 and Id-2 mRNA levels and a twofold increase in Id-3 mRNA levels in pancreatic cancer by comparison with the normal pancreas. In contrast, overall Id mRNA levels are not increased in CP.

Patients and Methods

Normal human pancreatic tissue samples from 7 male and 5 female donors (median age 41.8 years, range 14–68 years), CP tissues from 13 males and 1 female (median age 42.1 years; range 30–56 years), and pancreatic cancer tissues from 10 male and 6 female donors (median age 62.6 years; range 53–83 years) were obtained through an organ donor program and from surgical specimens from patients with severe symptomatic chronic pancreatitis or pancreatic cancer. A partial duodenopancreatectomy (Whipple/pylorus-preserving Whipple; $n = 13$), a left resection of the pancreas ($n = 2$), or a total pancreatectomy ($n = 1$) were carried out in the pancreatic cancer patients. According to the TNM classification of the Union Internationale Contre le Cancer (UICC) 6 tumors were stage 1, 1 was stage 2, and 9 were stage 3 ductal cell adenocarcinoma. Freshly removed tissue samples were fixed in 10% formaldehyde solution for 12 to 24 hours and paraffin-embedded for histological analysis. In addition, tissue samples were frozen in liquid nitrogen immediately on surgical removal and maintained in −80°C until use for RNA extraction. All studies were approved by the Ethics Committee of the University of Bern, Bern, Switzerland, and by the Human Subjects Committee at the University of California, Irvine, California.

Northern Blot Analysis

Northern blot analysis was carried out as described previously.[26,28] Briefly, total RNA was extracted by the single step acid guanidinium thiocyanate phenol chloroform method. RNA was size-fractionated on 1.2% agarose/1.8 mol/L formaldehyde gels, electrotransferred onto nylon membranes, and cross-linked by UV irradiation. Blots were prehybridized and hybridized with cDNA probes and washed under high stringency conditions. The following cDNA probes were used: a 979-bp human Id-1 cDNA probe, a 440-bp human Id-2 cDNA probe, and a 450-bp human Id-3 cDNA probe, covering the entire coding regions of Id-1, Id-2, and Id-3, respectively. A *Bam*HI 190-bp fragment of mouse 7S cDNA that hybridizes with human cytoplasmic RNA was used to confirm equal RNA loading and transfer. Blots were then exposed at −80°C to Kodak BioMax-MS films and the resulting autoradiographs were scanned to quantify the intensity of the radiographic bands.[26,28] For each sample the ratio of Id mRNA expression to 7S expression was calculated. To compare the relative increase in expression of the respective Id mRNA species in the cancer and CP samples, the same normal samples were used for normal/cancer and normal/CP membranes. The median score for Id-1, Id-2, and Id-3 mRNA levels in these normal samples was set to 100. Statistical analysis was performed with SigmaStat software (Jandel Scientific, San Raphael, CA). The rank sum test was used, and $P < 0.05$ was taken as the level of significance.

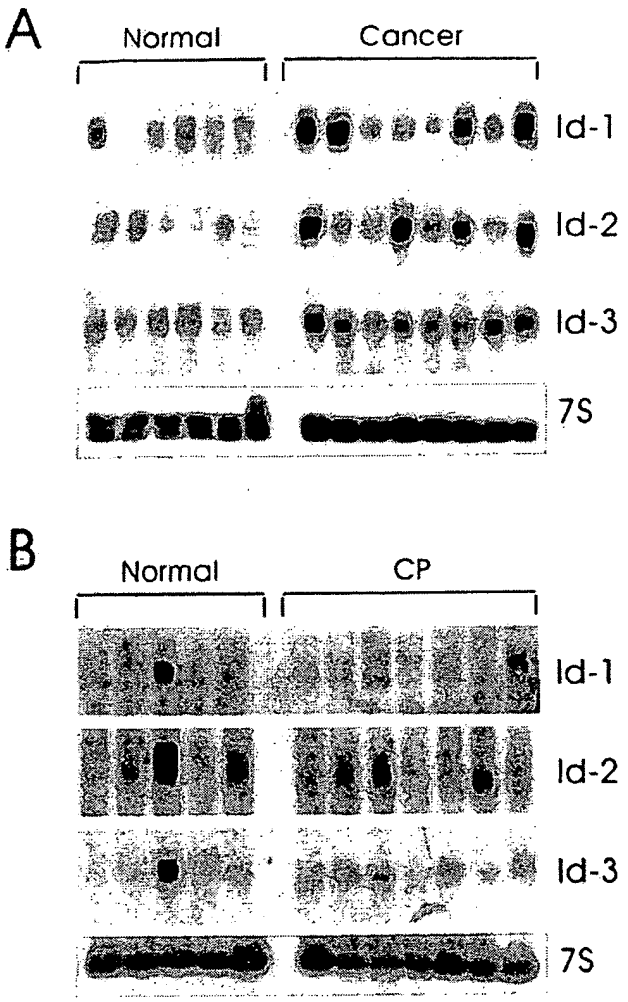

Figure 1. mRNA expression of Id-1, Id-2, and Id-3 in pancreatic cancer and chronic pancreatitis. Total RNA (20 μg/lane) from six normal, eight cancerous, and seven chronic pancreatitis tissue samples were subjected to Northern blot analysis using ^{32}P-labeled cDNA probes (500,000 cpm/ml) specific for Id-1, Id-2, and Id-3, respectively. A 7S cDNA probe (50,000 cpm/ml) was used as a loading and transfer control. Exposure times of the normal/cancer blots were 1 day for all Id probes, and 2 days for the normal/CP blots. Exposure time was 4 hours for mouse 7S cDNA. By comparison with the normal samples, Id-1 and Id-3 mRNA levels were elevated in 8 and 9 cancer samples, respectively, whereas Id-2 was elevated in 6 cancer samples.

Cell Culture and Western Blot Analysis

PANC-1, MIA-PaCa-2, ASPC-1, and CAPAN-1 human pancreatic cell lines were obtained from ATCC (Manassas, VA). COLO-357 human pancreatic cells were a gift from Dr. R. S. Metzger (Durham, NC). Cells were routinely grown in DMEM (COLO-357, MIA-PaCa-2, PANC-1) or RPMI (ASPC-1, CAPAN-1) supplemented with 10% fetal bovine serum, 100 U/ml penicillin, and 100 μg/ml streptomycin. For immunoblot analysis, exponentially growing

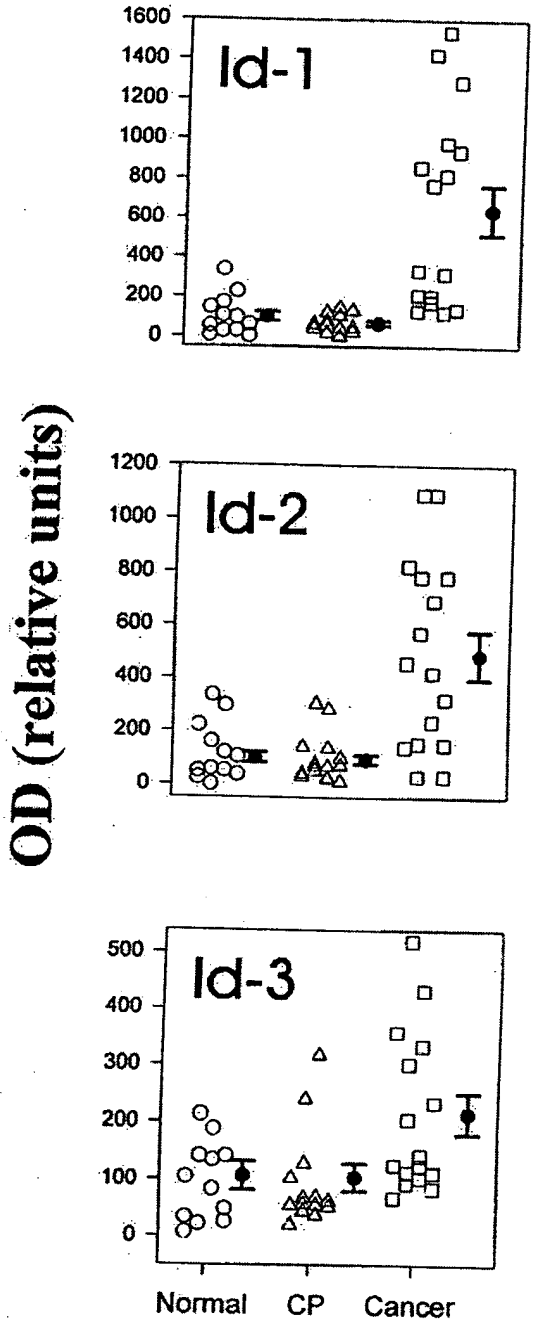

Figure 2. Densitometric analysis of Northern blots. Autoradiographs of Northern blots from 12 normal, 14 CP, and 16 pancreatic cancers were analyzed by densitometry. mRNA levels were determined by calculating the ratio of the optical density for the respective Id mRNA species in relation to the optical density of mouse 7S cDNA. To compare the relative increase in expression of the respective Id mRNA species in the cancer and CP samples, the same normal samples were used for normal/cancer and normal/CP membranes. Normal pancreatic tissues are indicated by circles, CP tissues by triangles, and cancer tissues by squares. Data are expressed as median scores ± SD. By comparison with the normal samples, only the cancer samples exhibited significant increases: 6.5-fold ($P < 0.01$) for Id-1, fivefold ($P < 0.01$) for Id-2, and twofold ($P = 0.027$) for Id-3.

Figure 3. Id mRNA and protein expression in pancreatic cancer cell lines. Upper panels: Total RNA (20 μg/lane) from 5 pancreatic cancer cell lines were subjected to Northern blot analysis using ^{32}P-labeled cDNA probes (500,000 cpm/ml) specific for Id-1, Id-2, and Id-3, respectively. Exposure times were 1 day for all Id probes. Lower panels: Immunoblotting. Cell lysates (30 μg/lane) were subjected to SDS-PAGE. Membranes were probed with specific Id-1, Id-2, and Id-3 antibodies. Visualization was performed by enhanced chemiluminescence.

cells (60–70% confluent) were solubilized in lysis buffer containing 50 mmol/L Tris-HCl, pH 7.4, 150 mmol/L NaCl, 1 mmol/L EDTA, 1 μg/ml pepstatin A, 1 mmol/L phenylmethylsulfonyl fluoride (PMSF), and 1% Triton X-100. Proteins were subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), transferred to Immobilon P membranes, and incubated for 90 minutes with the indicated antibodies and for 60 minutes with secondary antibodies against rabbit IgG. Visualization was performed by enhanced chemiluminescence.

Immunohistochemistry

Specific rabbit anti-human Id-1 (C-20), Id-2 (C-20), and Id-3 (C-20; all from Santa Cruz Biotechnology, Santa

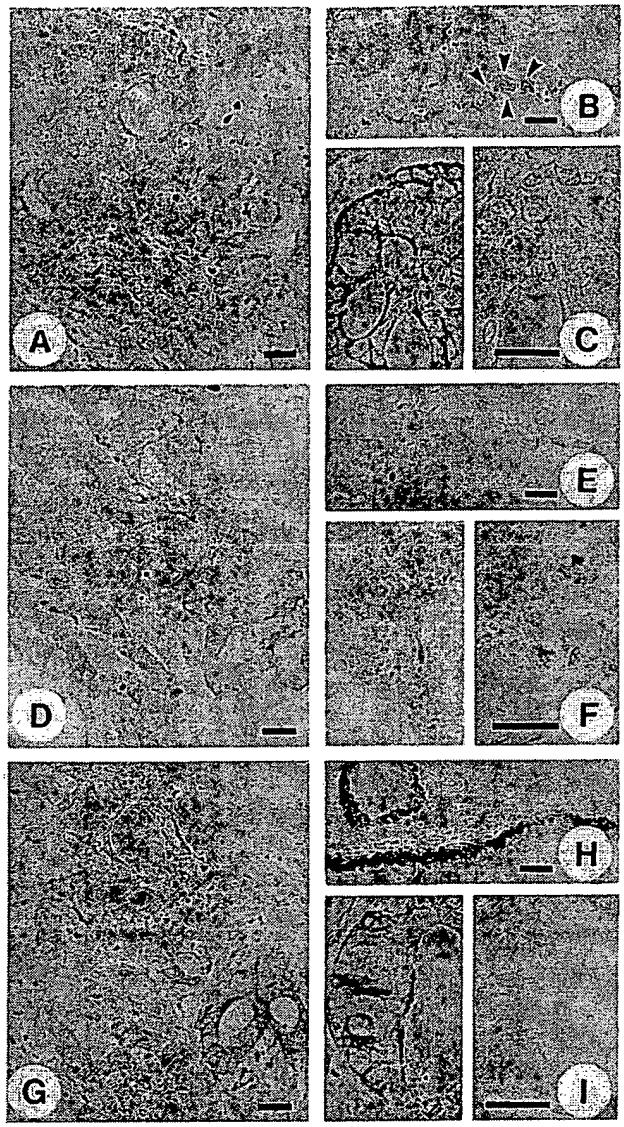

Figure 4. Normal and cancerous pancreatic tissues were subjected to immunostaining using highly specific anti-Id-1 (A-C), anti-Id-2 (D-F), and anti-Id-3 (G-I) antibodies as described in the Methods section. Moderate to strong Id-1 immunoreactivity was present in the cytoplasm of duct-like cancer cells (A and C, left panel). In the normal pancreas there was weak Id-1 immunoreactivity in the ductal cells (B). Preabsorption with the Id-1-specific blocking peptide abolished the Id-1 immunoreactivity (C, right panel). Strong Id-2 immunoreactivity was observed in the cytoplasm of the cancer cells that exhibited duct-like structures (D and F, left panel), whereas in the normal pancreas, there was only weak Id-2 immunoreactivity in the ductal cells (E). Preabsorption with the Id-2-specific blocking peptide abolished the Id-2 immunoreactivity (F, right panel). Moderate to strong Id-3 immunoreactivity was present in the duct-like cancer cells (G and I, left panel). Moderate to strong Id-3 immunoreactivity was also present in the ductal cells of normal pancreatic tissue samples (H). Id-3 immunoreactivity was completely abolished by preabsorption with the Id-3 specific blocking peptide (I, right panel). A, D, and G constitute serial sections of a pancreatic cancer sample, revealing coexpression of the three Id proteins. Scale bars, 25 μm.

Cruz, CA) polyclonal antibodies were used for immunhistochemistry. These affinity-purified rabbit polyclonal antibodies specifically react with Id-1, Id-2, and Id-3, respectively, of human origin, as determined by Western blotting. Paraffin-embedded sections (4 μm) were subjected to immunostaining using the streptavidin-peroxidase technique. Where indicated, immunostaining for all three Id proteins was performed on serial sections. Endogenous peroxidase activity was blocked by incubation for 30 minutes with 0.3% hydrogen peroxide in methanol. Tissue sections were incubated for 15 minutes (23°C) with 10% normal goat serum and then incubated for 16 hours at 4°C with the indicated antibodies in PBS containing 1% bovine serum albumin. Bound antibodies were detected with biotinylated goat anti-rabbit IgG secondary antibodies and streptavidin-peroxidase complex, using diaminobenzidine tetrahydrochloride as the substrate. Sections were counterstained with Mayer's hematoxylin. Preabsorption with Id-1-, Id-2-, or Id-3-specific blocking peptides completely abolished immunoreactivity of the respective primary antibody. The immunohistochemical results were semiquantitatively analyzed as described previously.[29,30] The percentage of positive cancer cells was stratified into four groups: 0, no cancer cells exhibiting immunoreactivity; 1, <33% of the cancer cells exhibiting immunoreactivity; 2, 33 to 67% of the cancer cells exhibiting immunoreactivity; 3 >67% of the cancer cells exhibiting immunoreactivity. The intensity of the immunohistochemical signal was also stratified into four groups: 0, no immunoreactivity; 1, weak immunoreactivity; 2, moderate immunoreactivity; 3, strong immunoreactivity. Finally, the sum of the results of the cell score and the intensity score was calculated. Statistical analysis was performed with SigmaStat software. The rank sum test was used, and $P < 0.05$ was taken as the level of significance.

Results

Northern blot analysis of total RNA isolated from 12 normal pancreatic tissues and 16 pancreatic cancers revealed the presence of the 1.2-kb Id-1 transcript and the 1.6-kb Id2 mRNA transcript in 11 of the 12 normal pancreatic samples, and the 1.3-kb Id-3 mRNA transcript in all normal pancreatic samples (Figure 1A, 2). In the cancer tissues, Id-1 mRNA levels were elevated in 8 of 16 samples, Id-2 mRNA levels were elevated in 9 of these samples, and Id-3 mRNA levels were elevated in 6 of these samples (Figure 1A, 2). Concomitant overexpression of all three Id species was observed in 6 of the cancer samples (38%). In contrast, none of the Id mRNA species were overexpressed in CP by comparison with normal controls (Figure 1B, 2). Densitometric analysis of all of the autoradiograms indicated that there was a 6.5-fold increase ($P < 0.01$) in Id-1 mRNA levels, a fivefold increase ($P < 0.01$) in Id-2 mRNA levels, and a twofold increase ($P = 0.027$) in Id-3 mRNA levels in the pancreatic cancer samples in comparison to normal controls (Figure 2). In contrast, there was no statistically significant difference in the expression levels of Id-1, Id-2, and Id-3, in CP tissues in comparison to the corresponding levels in the normal pancreas (Figure 2).

Next, we assessed the expression of the three Id genes in 5 human pancreatic cancer cell lines by Northern and Western blot analyses. Id-1 mRNA was present at varying levels in all 5 cell lines (Figure 3). ASPC-1, CAPAN-1, MIA-PaCa-2, and PANC-1 expressed moderate to high levels of Id-1 mRNA, whereas COLO-357 cells

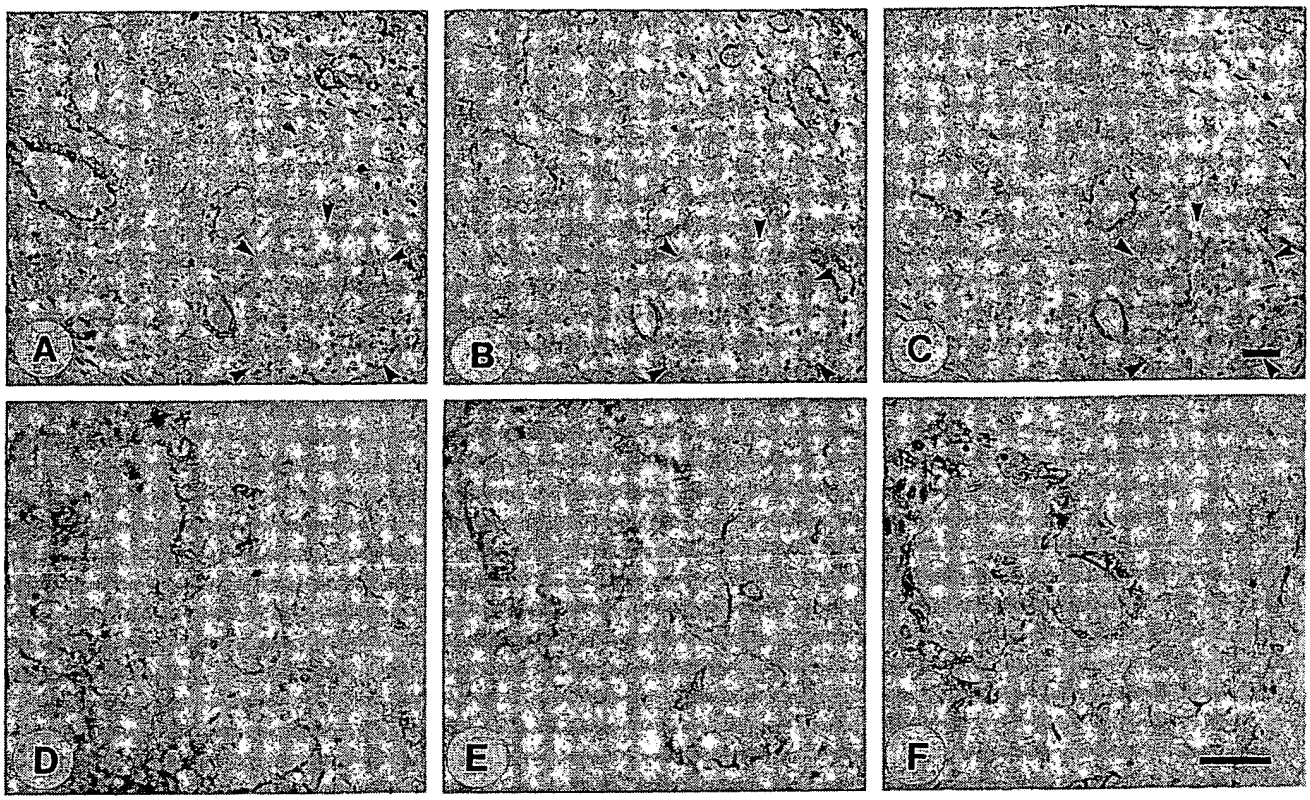

Figure 5. Immunohistochemistry of pancreatic cancer and dysplastic ducts in CP tissues. In the pancreatic cancer tissues (A-C) there was moderate to strong Id-1 (A), Id-2 (B), and Id-3 (C) immunoreactivity in the ductal cells in the areas adjacent to the cancer cells that exhibited CP-like alterations. Islet cells did not exhibit Id immunoreactivity (outlined by solid arrowheads). In the CP samples, moderate to strong Id-1 (D), Id-2 (E), and Id-3 (F) immunoreactivity was present in the cytoplasm of epithelial cells forming large dysplastic ducts. Scale bar, 25 μm.

expressed relatively low levels of this mRNA moiety. Western blotting with a highly specific anti-Id-1 antibody confirmed the presence of the approximately 14-kd Id-1 protein in the 4 cell lines that expressed high levels of Id-1 mRNA (Figure 3). Furthermore, the three cell lines with the highest Id-1 mRNA expression (CAPAN-1, MIA-PaCa-2, and PANC-1) also exhibited the highest Id-1 protein expression. Variable levels of the 1.6-kb Id-2 mRNA transcript were present in all 5 cell lines. In addition, a minor band of approximately 1.2 kb was visible in COLO-357 and MIA-PaCa-2 cells. Immunoblot analysis with a highly specific anti-Id-2 antibody revealed two bands of approximately 16 and 18 kd at relatively high levels in all of the cell lines with exception of PANC-1 cells, in which the 16-kd band was relatively faint (Figure 3). With the exception of MIA-PaCa-2 cells, there was a good correlation between Id-2 mRNA and protein levels (Figure 3). Id-3 mRNA was present at high levels in MIA-PaCa-2 cells, at moderate levels in COLO-357 cells, and at low levels in PANC-1 cells. Id-3 mRNA was not detectable in ASPC-1 and CAPAN-1 cells (Figure 3). Immunoblot analysis with a highly specific anti-Id-3 antibody revealed an approximately 14-kd band that was most abundant in MIA-PaCa-2 cells, and was also readily apparent in COLO-357 and PANC-1 cells. In contrast, only a faint Id-3 band was seen in ASPC-1 and CAPAN-1 cells. Thus, with the exception of PANC-1 cells, there was a good correlation between Id-3 mRNA and protein levels.

To determine the localization of Id-1, Id-2, and Id-3, immunostaining was carried out using the same highly specific anti-Id antibodies. In the pancreatic cancers, moderate to strong Id-1 immunoreactivity was present in the cancer cells in 9 of 10 randomly selected cancer samples. An example of moderate Id-1 immunoreactivity is shown in Figure 4A, and of strong immunoreactivity in Figure 4C (left panel). In contrast, in the normal pancreas, faint Id-1 immunoreactivity was present only in the ductal cells of pancreatic ducts (Figure 4B, arrowheads). Preabsorption with the Id-1-specific blocking peptide completely abolished the Id-1 immunoreactivity (Figure 4C, right panel). The cancer cells also exhibited strong Id-2 (Figure 4, D and F, left panel) and moderate to strong Id-3 immunoreactivity. An example of moderate Id-3 immunoreactivity is shown in Figure 4G, and of strong immunoreactivity in Figure 4I (left panel). In contrast, only faint Id-2 immunoreactivity was present in the ductal cells in the normal pancreas (Figure 4E), whereas Id-3 immunoreactivity in these cells was more variable and ranged from moderate to occasionally strong (Figure 4H). Islet cells and acinar cells were always devoid of Id immunoreactivity. Preabsorption of the respective antibody with the blocking peptides specific for Id-2 (Figure 4F, right panel) and Id-3 (Figure 4I, right panel) completely abolished immunoreactivity. Analysis of serial pancreatic cancer sections revealed that there was often colocalization of the

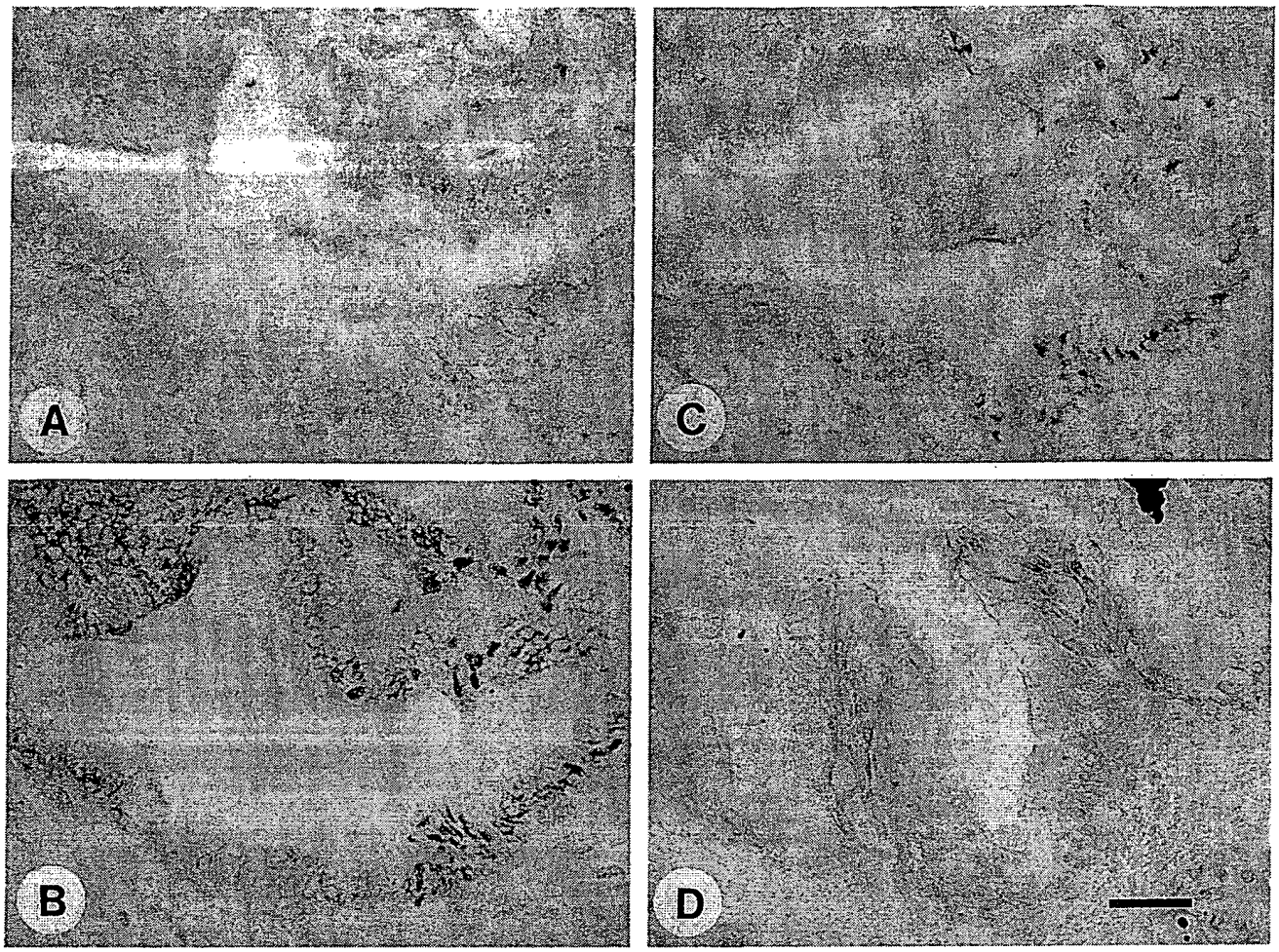

Figure 6. Immunohistochemistry of atypical papillary epithelium in CP tissues. Serial section analysis of some CP samples revealed the presence of large duct-like structures with atypical papillary epithelium. Mild to moderate Id-1 (A) and Id-2 (B) immunoreactivity and weak Id-3 (C) immunoreactivity was present in the cytoplasm of the cells forming these large ducts with papillary structures. Some CP samples also exhibited moderate Id-3 immunoreactivity in these cells (D). Scale bar, 25 μm.

three Id proteins. An example of serial sections from a pancreatic cancer tissue is shown in Figure 4, A, D, and G.

Id-1, Id-2, and Id-3 immunoreactivity was also present at moderate levels in the cytoplasm of ductal cells within CP-like areas adjacent to the cancer cells (Figure 5, A-C). As in the normal pancreas, islet cells (outlined by arrowheads) did not exhibit Id immunoreactivity. In 4 of 9 CP samples, there were foci of ductal cell dysplasia of relatively large interlobular ducts, all of which exhibited moderate to strong Id-1, Id-2, and Id-3 immunoreactivity (Figure 5, D-F). Five of 9 CP samples also contained foci of large ducts exhibiting atypical papillary epithelium. Serial section analysis of one of those CP samples revealed mild to moderate Id-1 and Id-2 immunoreactivity and weak Id-3 immunoreactivity in the cells of these atypical papillary ducts (Figure 6, A-C). In contrast, in some of these CP samples, moderate to strong Id-3 immunoreactivity was also observed (Figure 6D). However, most of the ductal cells forming the typical ductular structures of CP, such as large interlobular ducts and small proliferating ducts, exhibited generally only weak to occasionally moderate Id immunoreactivity (data not shown).

The immunohistochemical data for Id-1, Id-2, and Id-3 are summarized in Table 1. In the case of Id-1 and Id-2, the cancer cells as well as the dysplastic and atypical papillary ducts in CP exhibited a significantly higher score than the ductal cells in the normal pancreas. In contrast, due to the marked variability in Id-3 immunostaining in the normal pancreas, the differences between normal and cancer cells and normal and dysplastic cells did not achieve statistical significance.

Discussion

Id proteins constitute a family of HLH transcription factors that are important regulators of cellular differentiation and proliferation.[1,2] To date, four members of the human Id family have been identified.[1,10–12] Their expression is enhanced during cellular proliferation and in response to mitogenic stimuli,[19,31] and overexpression of Id genes inhibits differentiation and/or enhances proliferation in several different cell types.[15,32–34] The forced expression of Id-1 in mouse small intestinal epithelium results in

Table 1. Histological Scoring

		Id-1	Id-2	Id-3
Normal (n = 6)	Ductal cells	2.0 ± 0.4	2.3 ± 0.2	2.5 ± 0.9
Cancer (n = 10)	Cancer cells	4.5* ± 0.5	5.2§ ± 0.3	4.5 ± 0.6
CP (n = 9)	Typical CP lesions (n = 9)	2.7 ± 0.5	3.1 ± 0.6	3.4 ± 0.7
	Dysplastic ducts (n = 4)	5.3† ± 0.2	5.8‡ ± 0.2	5.3 ± 0.4
	Atypical papillary ducts (n = 5)	4.4‡ ± 0.2	5.2‡ ± 0.2	5.0 ± 0.4

Scoring of the histological specimens was performed as described in the Patients and Methods section. Values are the means ± SD of the number of samples indicated in parenthesis. P values are based on comparisons with the respective controls in the normal samples.
*, $P < 0.02$; †$P < 0.01$; ‡$P = 0.004$; §$P = 0.001$.

adenoma formation in these animals.[35] The growth-promoting effects of Id genes are thought to occur through several mechanisms. For example, Id-2 can bind to members of the pRB tumor suppressor family, thus blocking their growth-suppressing activity,[20,21] and Id-1 and Id-2 can antagonize the bHLH-mediated activation of known inhibitors of cell cycle progression such as the cyclin-dependent kinase inhibitor p21.[23]

In the present study, we determined by Northern blot analysis that a significant percentage of human pancreatic cancers expressed increased Id-1, Id-2, and Id-3 mRNA levels. Increased expression was most evident for Id-1 (6.5-fold) and Id-2 (fivefold). In contrast, Id-3 mRNA levels were only twofold increased in the cancer samples, partly because this mRNA was present at relatively high levels in the normal pancreas. Immunhistochemical analysis confirmed the presence of Id-1, Id-2, and Id-3 in the cancer cells within the tumor mass, whereas in the normal pancreas faint Id-1 and Id-2 immunoreactivity and moderate to occasionally strong Id-3 immunoreactivity was present in some ductal cells. Pancreatic acinar and islet cells in the normal pancreas were devoid of Id-1, Id-2, and Id-3 immunoreactivity. In the cancer samples, all three Id proteins often colocalized in the cancer cells. Coexpression of all three Id genes was also observed in cultured pancreatic cancer cell lines, which often exhibited a close correlation between Id mRNA and protein expression. However, in MIA-PaCa-2 there was a divergence of Id-2 mRNA and protein levels, and in PANC-1 cells, Id-3 mRNA levels did not correlate well with Id-3 protein expression. These observations suggest that in these cells, the half-life of either Id mRNA or Id protein may be altered by comparison with the other cell lines. Interestingly, Id-2 immunoblotting revealed two closely spaced bands of approximately 16 and 18 kd in 4 of 5 cell lines. In view of the fact that two possible initiation codons have been reported for the Id-2 gene,[36] our observation raises the possibility that the two Id-2-immunoreactive bands may represent separate translation products of the Id-2 gene.

Pancreatic cancers often harbor p53 tumor suppressor gene mutations[37] and exhibit alterations in apoptosis pathways. Thus, these cancers often exhibit increased expression of anti-apoptotic proteins such as Bcl-2[38] and abnormal resistance to Fas-ligand-mediated apoptosis.[39] It has been shown recently that forced constitutive expression of Id genes together with the expression of anti-apoptotic genes such as Bcl-2 or $BclX_L$ can result in malignant transformation of human fibroblasts,[11] raising the possibility that the enhanced Id expression in pancreatic cancers together with increased expression of anti-apoptotic genes may contribute to the malignant potential of pancreatic cancer cells *in vivo*.

In the CP tissues there was no significant increase in Id-1, Id-2, and Id-3 mRNA levels in comparison to the normal pancreas. Immunohistochemical analysis of pancreatic cancer samples revealed colocalization of weak to moderate Id-1, Id-2, and Id-3 immunoreactivity in proliferating ductal cells in the CP-like regions adjacent to the cancer cells, indicating that Id expression was not restricted to the cancer cells. Similarly, analysis of CP samples indicated weak Id-1, Id-2, and Id-3 immunoreactivity in the cells of small proliferating ducts and large ducts without dysplastic changes. In general, there was a correlation between weak immunoreactivity and low Id mRNA levels. However, in samples that harbored large ducts with papillary structures there was moderate Id immunoreactivity, and in the cells forming dysplastic ducts there was moderate to strong Id immunoreactivity. In these CP samples, Id mRNA levels were relatively higher than in the CP samples that were devoid of these histological changes. Overall, however, increased Id expression, most notably of Id-1 and Id-2, distinguished a subgroup of pancreatic cancers from CP (Table 1).

Epidemiological studies have shown that the risk of developing pancreatic cancer is increased up to 16-fold in patients with pre-existing CP in comparison to the general population.[40] The mechanisms that contribute to neoplastic transformation in CP are not known. Although there is no established tumor progression model for pancreatic cancer, such as the adenoma-carcinoma sequence of colorectal carcinoma,[41] it is generally accepted that K-*ras* and p16 mutations occur relatively early in pancreatic carcinogenesis, whereas p53 mutations occur late in this process.[37,41–43] Increased Id expression may contribute to malignant transformation of cultured cell lines *in vitro*[11] and has been linked to cell invasion in a murine mammary epithelial cell line.[44] In view of the current findings that Id-1, Id-2, and Id-3 are overexpressed in pancreatic cancer and in dysplastic/metaplastic ducts in CP, these observations raise the possibility that elevated levels of Id-1, Id-2, and, to a lesser extent, Id-3 may represent relatively early markers of pancreatic malignant transformation and may contribute to the pathobiology of pancreatic cancer.

References

1. Jan YN, Jan LY: HLH proteins, fly neurogenesis, and vertebrate myogenesis. Cell 1993, 75:827–830
2. Olson EN, Klein WH: bHLH factors in muscle development: dead lines and commitments, what to leave in and what to leave out. Genes Dev 1994, 8:1–8
3. Begley CG, Aplan PD, Denning SM, Haynes BF, Waldmann TA, Kirsch IR: The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. Proc Natl Acad Sci USA 1989, 86:10128–10132
4. Johnson JE, Birren SJ, Anderson DJ: Two rat homologues of Drosophila achaete-scute specifically expressed in neuronal precursors. Nature 1990, 346:858–861
5. Weintraub H: The MyoD family and myogenesis: redundancy, networks, and thresholds. Cell 1993, 75:1241–1244
6. Hu JS, Olson EN, Kingston RE: HEB, a helix-loop-helix protein related to E2A, and ITF2 that can modulate the DNA-binding ability of myogenic regulatory factors. Mol Cell Biol 1992, 12:1031–1042
7. Langlands K, Yin X, Anand G, Prochownik EV: Differential interactions of Id proteins with basic-helix-loop-helix transcription factors. J Biol Chem 1997, 272:19785–19793
8. Murre C, Bain G, van Dijk MA, Engel I, Furnari BA, Massari ME, Matthews JR, Quong MW, Rivera RR, Stuiver MH: Structure and function of helix-loop-helix proteins. Biochim Biophys Acta 1994, 1218:129–135
9. Murre C, McCaw PS, Vaessin H, Caudy M, Jan LY, Jan YN, Cabrera CV, Buskin JN, Hauschka SD, Lassar AB, Baltimore D: Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell 1989, 58: 537–544
10. Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H: The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell 1990, 61:49–59
11. Norton JD, Atherton GT: Coupling of cell growth control and apoptosis functions of Id proteins. Mol Cell Biol 1998, 18:2371–2381
12. Norton JD, Deed RW, Craggs G, Sablitzky F: Id helix-loop-helix proteins in cell growth and differentiation. Trends Cell Biol 1998, 8:58–65
13. Christy BA, Sanders LK, Lau LF, Copeland NG, Jenkins NA, Nathans D: An Id-related helix-loop-helix protein encoded by a growth factor-inducible gene. Proc Natl Acad Sci USA 1991, 88:1815–1819
14. Kawaguchi N, DeLuca HF, Noda M: Id gene expression and its suppression by 1,25-dihydroxyvitamin D3 in rat osteoblastic osteosarcoma cells. Proc Natl Acad Sci USA 1992, 89:4569–4572
15. Kreider BL, Benezra R, Rovera G, Kadesch T: Inhibition of myeloid differentiation by the helix-loop-helix protein Id. Science 1992, 255: 1700–1702
16. Le Jossic C, Ilyin GP, Loyer P, Glaise D, Cariou S, Guguen-Guillouzo C: Expression of helix-loop-helix factor Id-1 is dependent on the hepatocyte proliferation and differentiation status in rat liver and in primary culture. Cancer Res 1994, 54:6065–6068
17. Sun XH, Copeland NG, Jenkins NA, Baltimore D: Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. Mol Cell Biol 1991, 11:5603–5611
18. Wilson RB, Kiledjian M, Shen CP, Benezra R, Zwollo P, Dymecki SM, Desiderio SV, Kadesch T: Repression of immunoglobulin enhancers by the helix-loop-helix protein Id: implications for B-lymphoid-cell development. Mol Cell Biol 1991, 11:6185–6191
19. Hara E, Yamaguchi T, Nojima H, Ide T, Campisi J, Okayama H, Oda K: Id-related genes encoding helix-loop-helix proteins are required for G1 progression and are repressed in senescent human fibroblasts. J Biol Chem 1994, 269:2139–2145
20. Iavarone A, Garg P, Lasorella A, Hsu J, Israel MA: The helix-loop-helix protein Id-2 enhances cell proliferation and binds to the retinoblastoma protein. Genes Dev 1994, 8:1270–1284
21. Lasorella A, Iavarone A, Israel MA: Id2 specifically alters regulation of the cell cycle by tumor suppressor proteins. Mol Cell Biol 1996, 16:2570–2578
22. Peverali FA, Ramqvist T, Saffrich R, Pepperkok R, Barone MV, Philipson L: Regulation of G1 progression by E2A and Id helix-loop-helix proteins. EMBO J 1994, 13:4291–4301
23. Prabhu S, Ignatova A, Park ST, Sun XH: Regulation of the expression of cyclin-dependent kinase inhibitor p21 by E2A and Id proteins. Mol Cell Biol 1997, 17:5888–5896
24. Warshaw AL, Fernandez-del Castillo C: Pancreatic carcinoma. N Engl J Med 1992, 326:455–465
25. Korc M: Role of growth factors in pancreatic cancer. Surg Oncol Clin North Am 1998, 7:25–41
26. Kleeff J, Ishiwata T, Friess H, Büchler MW, Israel MA, Korc M: The helix-loop-helix protein Id2 is overexpressed in human pancreatic cancer. Cancer Res 1998, 58:3769–3772
27. Oertel JE, Heffes CS, Oertel YC: Pancreas. Diagnostic Surgical Pathology. Edited by SS Sternberg. New York, Raven Press, 1989, pp 1057–1093
28. Korc M, Chandrasekar B, Yamanaka Y, Friess H, Büchler MW, Beger HG: Overexpression of the epidermal growth factor receptor in human pancreatic cancer is associated with concomitant increase in the levels of epidermal growth factor and transforming growth factor α. J Clin Invest 1992, 90:1352–1360
29. Saeki T, Stromberg K, Qi CF, Gullick WJ, Tahara E, Normanno N, Ciardiello F, Kenney N, Johmson GR, Salomon DS: Differential immunohistochemical detection of amphiregulin and cripto in human normal colon and colorectal tumors. Cancer Res 1992, 52:3467–3473
30. Cantero D, Friess H, Deflorin J, Zimmermann A, Bründler MA, Riesle E, Korc M, Büchler MW: Enhanced expression of urokinase plasminogen activator and its receptor in pancreatic carcinoma. Br J Cancer 1997, 75:388–395
31. Desprez PY, Hara E, Bissell MJ, Campisi J: Suppression of mammary epithelial cell differentiation by the helix-loop-helix protein Id-1. Mol Cell Biol 1995, 15:3398–3404
32. Shoji W, Yamamoto T, Obinata M: The helix-loop-helix protein Id inhibits differentiation of murine erythroleukemia cells. J Biol Chem 1994, 269:5078–5084
33. Cross JC, Flannery ML, Blanar MA, Steingrimsson E, Jenkins NA, Copeland NG, Rutter WJ, Werb Z: Hxt encodes a basic helix-loop-helix transcription factor that regulates trophoblast cell development. Development 1995, 121:2513–2523
34. Sun XH: Constitutive expression of the Id1 gene impairs mouse B cell development. Cell 1994, 79:893–900
35. Wice BM, Gordon JI: Forced expression of Id-1 in the adult mouse small intestinal epithelium is associated with development of adenomas. J Biol Chem 1998, 273:25310–25319
36. Barone MV, Pepperkok R, Peverali FA, Philipson L: Id proteins control growth induction in mammalian cells. Proc Natl Acad Sci USA 1994, 91:4985–4988
37. Barton CM, Staddon SL, Hughes CM, Hall, PA, O'Sullivan C, Kloppel G, Theis B, Russell RC, Neoptolmos J, Williamson RCN, Lane DP, Lemoine NR: Abnormalities of the p53 tumour suppressor gene in human pancreatic cancer. Br J Cancer 1991, 64:1076–1082
38. Ohshio G, Suwa H, Imamura T, Yamaki K, Tanaka T, Hashimoto Y, Imamura M: An immunohistochemical study of bcl-2 and p53 protein expression in pancreatic carcinomas. Scand J Gastroenterol 1998, 33:535–539
39. Ungefroren H, Voss M, Jansen M, Roeder C, Henne-Bruns D, Kremer B, Kalthoff H: Human pancreatic adenocarcinomas express Fas and Fas ligand yet are resistant to Fas-mediated apoptosis. Cancer Res 1998, 58:1741–1749
40. Niederau C, Niederau MC, Heintges T, Lüthen R: Epidemiology: relation between chronic pancreaitis and pancreatic carcinoma. Cancer of the Pancreas. Edited by HG Beger, MW Büchler, MH Schoenberg. Ulm, Germany, Universitätsverlag Ulm GmbH, 1996, pp 6–9
41. Moskaluk CA, Kern SE: Molecular genetics of pancreatic carcinoma. Pancreatic Cancer: Pathogenesis, Diagnosis, and Treatment. Edited by HA Reber. Totowa, NJ, Humana Press, 1998, pp 3–20
42. Moskaluk CA, Hruban RH, Kern SE: p16 and K-ras gene mutations in the intraductal precursors of human pancreatic adenocarcinoma. Cancer Res 1997, 57:2140–2143
43. Tada M, Ohashi M, Shiratori Y, Okudaira T, Komatsu Y, Kawabe T, Yoshida H, Machinami R, Kishi K, Omata M: Analysis of K-ras gene mutation in hyperplastic duct cells of the pancreas without pancreatic disease. Gastroenterology 1996, 110:227–231
44. Desprez PY, Lin CQ, Thomasset N, Sympson CJ, Bissell MJ, Campisi J: A novel pathway for mammary epithelial cell invasion induced by the helix-loop-helix protein Id-1. Mol Cell Biol 1998, 18:4577–4588

MOLECULAR AND CELLULAR BIOLOGY, Nov. 1999, p. 7357–7368 Vol. 19, No. 11
0270-7306/99/$04.00+0

A Sampling of the Yeast Proteome

B. FUTCHER,[1]* G. I. LATTER,[1] P. MONARDO,[1] C. S. MCLAUGHLIN,[2] AND J. I. GARRELS[3]

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724[1]; Department of Biological Chemistry, University of California, Irvine, California 92717[2]; and Proteome, Inc., Beverly, Massachusetts 01915[3]

Received 15 June 1999/Returned for modification 16 July 1999/Accepted 28 July 1999

In this study, we examined yeast proteins by two-dimensional (2D) gel electrophoresis and gathered quantitative information from about 1,400 spots. We found that there is an enormous range of protein abundance and, for identified spots, a good correlation between protein abundance, mRNA abundance, and codon bias. For each molecule of well-translated mRNA, there were about 4,000 molecules of protein. The relative abundance of proteins was measured in glucose and ethanol media. Protein turnover was examined and found to be insignificant for abundant proteins. Some phosphoproteins were identified. The behavior of proteins in differential centrifugation experiments was examined. Such experiments with 2D gels can give a global view of the yeast proteome.

The sequence of the yeast genome has been determined (9). More recently, the number of mRNA molecules for each expressed gene has been measured (27, 30). The next logical level of analysis is that of the expressed set of proteins. We have begun to analyze the yeast proteome by using two-dimensional (2D) gels.

2D gel electrophoresis separates proteins according to isoelectric point in one dimension and molecular weight in the other dimension (21), allowing resolution of thousands of proteins on a single gel. Although modern imaging and computing techniques can extract quantitative data for each of the spots in a 2D gel, there are only a few cases in which quantitative data have been gathered from 2D gels. 2D gel electrophoresis is almost unique in its ability to examine biological responses over thousands of proteins simultaneously and should therefore allow us a relatively comprehensive view of cellular metabolism.

We and others have worked toward assembling a yeast protein database consisting of a collection of identified spots in 2D gels and of data on each of these spots under various conditions (2, 7, 8, 10, 23, 25). These data could then be used in analyzing a protein or a metabolic process. *Saccharomyces cerevisiae* is a good organism for this approach since it has a well-understood physiology as well as a large number of mutants, and its genome has been sequenced. Given the sequence and the relative lack of introns in *S. cerevisiae*, it is easy to predict the sequence of the primary protein product of most genes. This aids tremendously in identifying these proteins on 2D gels.

There are three pillars on which such a database rests: (i) visualization of many protein spots simultaneously, (ii) quantification of the protein in each spot, and (iii) identification of the gene product for each spot. Our first efforts at visualization and identification for *S. cerevisiae* have been described elsewhere (7, 8). Here we describe quantitative data for these proteins under a variety of experimental conditions.

MATERIALS AND METHODS

Strains and media. *S. cerevisiae* W303 (*MATa ade2-1 his3-11,15 leu2-3, 112 trp1-1 ura3-1 can1-100*) was used (26). −Met YNB (yeast nitrogen base) medium was 1.7 g of YNB (Difco) per liter, 5 g of ammonium sulfate per liter, and adenine, uracil, and all amino acids except methionine; −Met −Cys YNB medium was the same but without methionine or cysteine. Medium was supplemented with 2% glucose (for most experiments) or with 2% ethanol (for ethanol experiments). Low-phosphate YEPD was described by Warner (28).

Isotopic labeling of yeast and preparation of cell extracts. Yeast strains were labeled and proteins were extracted as described by Garrels et al. (7, 8). Briefly, cells were grown to 5×10^6 cells per ml. at 30°C; 1 ml of culture was transferred to a fresh tube, and 0.3 mCi of [^{35}S]methionine (e.g., Express protein labeling mix; New England Nuclear) was added to this 1-ml culture. The cells were incubated for a further 10 to 15 min and then transferred to a 1.5-ml microcentrifuge tube, chilled on ice, and harvested by centrifugation. The supernatant was removed, and the cell pellet was resuspended in 100 μl of lysis buffer (20 mM Tris-HCl [pH 7.6], 10 mM NaF, 10 mM sodium pyrophosphate, 0.5 mM EDTA, 0.1% deoxycholate; just before use, phenylmethylsulfonyl fluoride was added to 1 mM, leupeptin was added to 1 μg/ml, pepstatin was added to 1 μg/ml, tosylsulfonyl phenylalanyl chloromethyl ketone was added to 10 μg/ml, and soybean trypsin inhibitor was added to 10 μg/ml).

The resuspended cells were transferred to a screw-cap 1.5-ml polypropylene tube containing 0.28 g of glass beads (0.5-mm diameter; Biospec Products) or 0.40 g of zirconia beads (0.5-mm diameter; Biospec Products). After the cap was secured, the tube was inserted into a MiniBeadbeater 8 (Biospec Products) and shaken at medium high speed at 4°C for 1 min. Breakage was typically 75%. Tubes were then spun in a microcentrifuge for 10 s at 5,000 × *g* at 4°C.

With a very fine pipette tip, liquid was withdrawn from the beads and transferred to a prechilled 1.5-ml tube containing 7 μl of DNase I (0.5 mg/ml; Cooper product no. 6330)–RNase A (0.25 mg/ml; Cooper product no. 5679)–Mg (50 mM $MgCl_2$) mix. Typically 70 μl of liquid was recovered. The mixture was incubated on ice for 10 min to allow the RNase and DNase to work.

Next, 75 μl of 2× dSDS (2× dSDS is 0.6% sodium dodecyl sulfate [SDS], 2% mercaptoethanol, and 0.1 M Tris-HCl [pH 8]) was added. The tube was plunged into boiling water, incubated for 1 min, and then plunged into ice. After cooling, the tube was centrifuged at 4°C for 3 min at 14,000 × *g*. The supernatant was transferred to a fresh tube and frozen at −70°C. About 5 μl of this supernatant was used for each 2D gel.

2D polyacrylamide gels. 2D gels were made and run as described elsewhere (6–8).

Image analysis of the gels. The Quest II software system was used for quantitative image analysis (20, 22). Two techniques were used to collect quantitative data for analysis by Quest II software. First, before the advent of phosphorimagers, gels were dried and fluorographed. Each gel was exposed to film for three different times (typically 1 day, 2 weeks, and 6 weeks) to increase the dynamic range of the data. The films were scanned along with calibration strips to relate film optical density to disintegrations per minute in the gels and analyzed by the software to obtain a linear relationship between disintegrations per minute in the spots and optical densities of the film images. The quantitative data are expressed as parts per million of the total cellular protein. This value is calculated from the disintegrations per minute of the sample loaded onto the gel and by comparing the film density of each data spot with density of the film over the calibration strips of known radioactivity exposed to the same film. This yields the disintegrations per minute per millimeter for each spot on the gel and thence its parts-per-minute value.

After the advent of phosphorimaging, gels bearing ^{35}S-labeled proteins were exposed to phosphorimager screens and scanned by a Fuji phosphorimager, typically for two exposures per gel. Calibration strips of known radioactivity were exposed simultaneously. Scan data from the phosphorimager was assimilated by Quest II software, and quantitative data were recorded for the spots on the gels.

* Corresponding author. Mailing address: Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724. Phone: (516) 367-8828. Fax: (516) 367–8369. E-mail: futcher@cshl.org.

Measurements of protein turnover. Cells in exponential phase were pulse-labeled with [^{35}S]methionine, excess cold Met and Cys were added, and samples of equal volume were taken from the culture at intervals up to 90 min (in one experiment) or up to 160 min (in a second experiment). Incorporation of ^{35}S into protein was essentially 100% by the first sample (10 min). Extracts were made, and equal fractions of the samples were loaded on 2D gels (i.e., the different samples had different amounts of protein but equal amounts of ^{35}S). Spots were quantitated with a phosphorimaging and Quest software.

The software was queried for spots whose radioactivity decreased through the time course. The algorithm examined all data points for all spots, drew a best-fit line through the data points, and looked for spots where this line had a statistically significant negative slope. In one of the experiments, there was one such spot. To the eye, this was a minor, unidentified spot seen only in the first two samples (10 and 20 min). In the other experiment, the Quest software found no spots meeting the criteria. Therefore, we concluded that none of the identified spots (and all but one of the visible spots) represented proteins with long half-lives.

Centrifugal fractionation. Cells were labeled, harvested, and broken with glass beads by the standard method described above except that no detergent (i.e., no deoxycholate) was present in the lysis buffer. The crude lysate was cleared of unbroken cells and large debris by centrifugation at 300 × *g* for 30 s. The supernatant of this centrifugation was then spun at 16,000 × *g* for 10 min to give the pellet used for Fig. 6B. The supernatant of the 16,000 × *g*, 10-min spin was then spun at 100,000 × *g* for 30 min to give the supernatant used for Fig. 6A.

Protein abundance calculations. A haploid yeast cell contains about 4×10^{-12} g of protein (1, 15). Assuming a mean protein mass of 50 kDa, there are about 50×10^6 molecules of protein per cell. There are about 1.8 methionines per 10 kDa of protein mass, which implies 4.5×10^8 molecules of methionine per cell (neglecting the small pool of free Met). We measured (i) the counts per minute in each spot on the 2D gels, (ii) the total number of counts on each gel (by integrating counts over the entire gel), and (iii) the total number of counts loaded on the gel (by scintillation counting of the original sample). Thus, we know what fraction of the total incorporated radioactivity is present in each spot. After correcting for the methionine (and cysteine [see below]) content of each protein, we calculated an absolute number of protein molecules based on the fraction of radioactivity in each spot and on 50×10^6 total molecules per cell.

The labeling mixture used contained about one-fifth as much radioactive cysteine as radioactive methionine. Therefore, the number of cysteine molecules per protein was also taken into account in calculating the number of molecules of protein, but Cys molecules were weighted one-fifth as heavily as Met molecules.

mRNA abundance calculations. For estimation of mRNA abundance, we used SAGE (serial analysis of gene expression) data (27) and Affymetrix chip hybridization data (29a, 30). The mRNA column in Table 1 shows mRNA abundance calculated from SAGE data alone. However, the SAGE data came from cells growing in YEPD medium, whereas our protein measurements were from cells growing in YNB medium. In addition, SAGE data for low-abundance mRNAs suffers from statistical variation. Therefore, we also used chip hybridization data (29a, 30) for mRNA from cells grown in YNB. These hybridization data also had disadvantages. First, the amounts of high-abundance mRNAs were systematically underestimated, probably because of saturation in the hybridizations, which used 10 μg of cRNA. For example, the abundance of *ADH1* mRNA was 197 copies per cell by SAGE but only 32 copies per cell by hybridization, and the abundance of *ENO2* mRNA was 248 copies per cell by SAGE but only 41 by hybridization. When the amount of cRNA used in the hybridization was reduced to 1 μg, the apparent amounts of mRNA were similar to the amounts determined by SAGE (29a, 29b). However, experiments using 1 μg of cRNA have been done for only some genes (29a). Because amounts of mRNA were normalized to 15,000 per cell, and because the amounts of abundant mRNAs were underestimated, there is a 2.2-fold overestimate of the abundance of nonabundant mRNAs. We calculated this factor of 2.2 by adding together the number of mRNA molecules from a large number of genes expressed at a low level for both SAGE data and hybridization data. The sum for the same genes from hybridization data is 2.2-fold greater than that from SAGE data.

To take into account these difficulties, we compiled a list of "adjusted" mRNA abundance as follows. For all high-abundance mRNAs of our identified proteins, we used SAGE data. For all of these particular mRNAs, chip hybridization suggested that mRNA abundance was the same in YEPD and YNB media. For medium-abundance mRNAs, SAGE data were used, but when hybridization data showed a significant difference between YEPD and YNB, then the SAGE data were adjusted by the appropriate factor. Finally, for low-abundance mRNAs, we used data from chip hybridizations from YNB medium but divided by 2.2 to normalize to the SAGE results. These calculations were completed without reference to protein abundance.

CAI. The codon adaptation index (CAI) was taken from the yeast proteome database (YPD) (13), for which calculations were made according to Sharp and Li (24). Briefly, the index uses a reference set of highly expressed genes to assign a value to each codon, and then a score for a gene is calculated from the frequency of use of the various codons in that gene (24).

Statistical analysis. The JMP program was used with the aid of T. Tully. The JMP program showed that neither mRNA nor protein abundances were normally distributed; therefore, Spearman rank correlation coefficients (r_s) were calculated. The mRNA (adjusted and unadjusted) and protein data were also transformed so that Pearson product-moment correlation coefficients (r_p) could be calculated. First, this was done by a Box-Cox transformation of log-transformed data. This transformation produced normal distributions, and an r_p of 0.76 was achieved. However, because the Box-Cox transformation is complex, we also did a simpler logarithmic transformation. This produced a normal distribution for the protein data. However, the distribution for the mRNA and adjusted mRNA data was close to, but not quite, normal. Nevertheless, we calculated the r_p and found that it was 0.76, identical to the coefficient from the Box-Cox transformed data. We therefore believe that this correlation coefficient is not misleading, despite the fact that the log(mRNA) distribution is not quite normal.

RESULTS

Visualization of 1,400 spots on three gel systems. Yeast proteins have isoelectric points ranging from 3.1 to 12.8, and masses ranging from less than 10 kDa to 470 kDa. It is difficult to examine all proteins on a single kind of gel, because a gel with the needed range in pI and mass would give poor resolution of the thousands of spots in the central region of the gel. Therefore, we have used three gel systems: (i) pH "4 to 8" with 10% polyacrylamide; (ii) pH "3 to 10" with 10% polyacrylamide; and (iii) nonequilibrium with 15% polyacrylamide (7, 8). Each gel system allows good resolution of a subset of yeast proteins.

Figure 1 shows a pH 4–8, 10% polyacrylamide gel. The pH at the basic end of the isoelectric focusing gel cannot be maintained throughout focusing, and so the proteins resolved on such gels have isoelectric points between pH 4 and pH 6.7. For these pH 4–8 gels, we see 600 to 900 spots on the best gels after multiple exposures.

The pH 3–10 gels (not shown) extend the pI range somewhat beyond pH 7.5, allowing detection of several hundred additional spots. Finally, we use nonequilibrium gels with 15% acrylamide in the second dimension. These allow visualization of about 100 very basic proteins and about 170 small proteins (less than 20 kDa). In total, using all three gel systems, about 1,400 spots can be seen. These represent about 1,200 different proteins, which is about one-quarter to one-third of the proteins expressed under these conditions (27, 30). Here, we focus on the proteins seen on the pH 4–8 gels.

Although nearly all expressed proteins are present on these gels, the number seen is limited by a problem we call coverage. Since there are thousands of proteins on each gel, many proteins comigrate or nearly comigrate. When two proteins are resolved, but are close together, and one protein spot is much more intense than the other, a problem arises in visualizing the weaker spot: at long exposures when the weak signal is strong enough for detection, the signal from the strong spot spreads and covers the signal from the weaker spot. Thus, weak spots can be seen only when they are well separated from strong spots.

For a given gel, the number of detectable spots initially rises with exposure time. However, beyond an optimal exposure, the number of distinguishable spots begins to decrease, because signals from strong spots cover signals from nearby weak spots. At long exposures, the whole autoradiogram turns black. Thus, there is an optimum exposure yielding the maximum number of spots, and at this exposure the weakest spots are not seen.

Largely because of the problem of coverage, the proteins seen are strongly biased toward abundant proteins. All identified proteins have a CAI of 0.18 or more, and we have identified no transcription factors or protein kinases, which are nonabundant proteins. Thus, this technology is useful for examining protein synthesis, amino acid metabolism, and glycolysis but not for examining transcription, DNA replication, or the cell cycle.

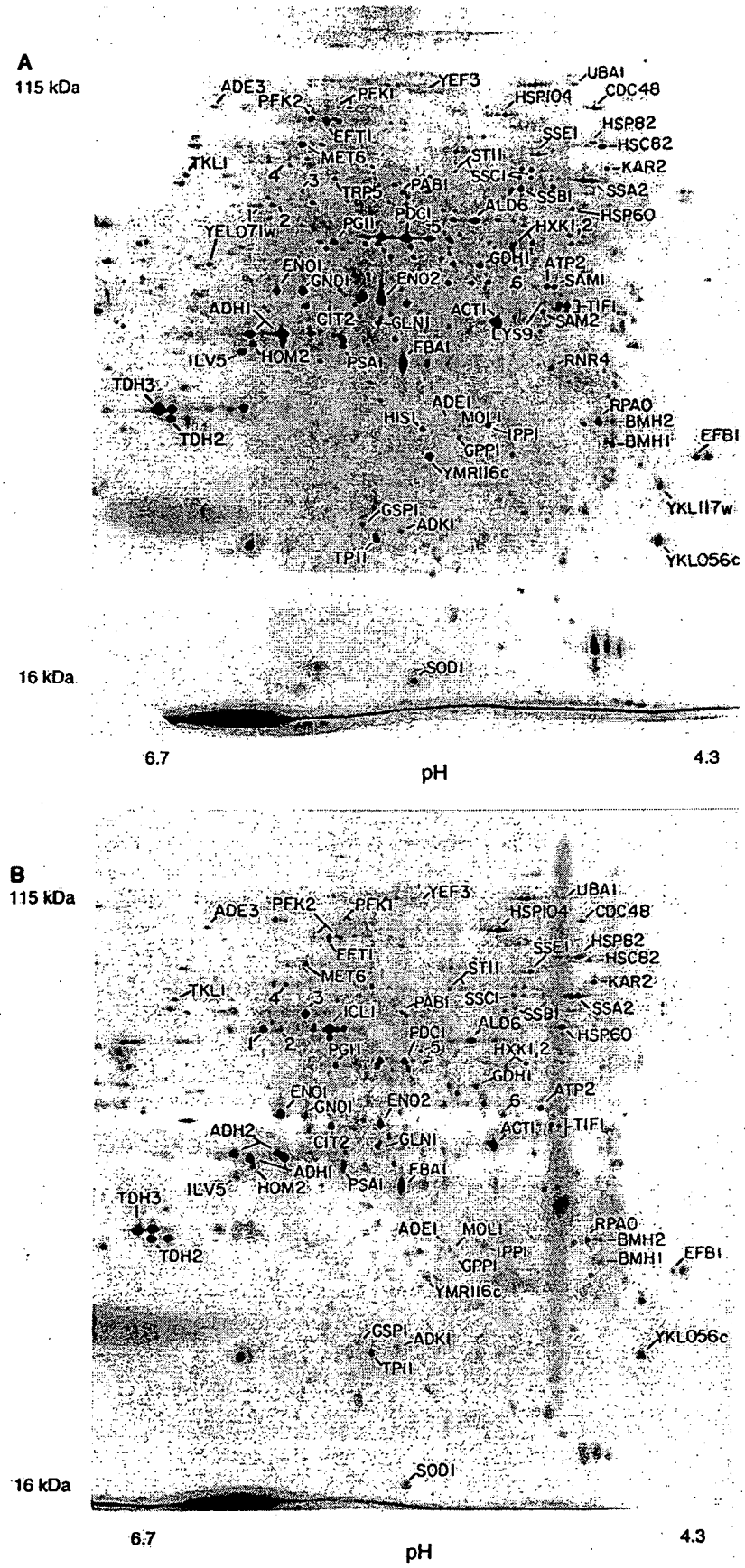

FIG. 1. 2D gels. The horizontal axis is the isoelectric focusing dimension, which stretches from pH 6.7 (left) to pH 4.3 (right). The vertical axis is the polyacrylamide gel dimension, which stretches from about 15 kDa (bottom) to at least 130 kDa (top). For panel A, extract was made from cells in log phase in glucose; for panel B, cells were grown in ethanol. The spots labeled 1 through 6 are unidentified proteins highly induced in ethanol.

Spot identification. The identification of various spots has been described elsewhere (7, 8). At present, 169 different spots representing 148 proteins have been identified. Many of these spots have been independently identified (2, 10, 23, 25). The main methods used in spot identification have been analysis of amino acid composition, gene overexpression, peptide sequencing, and mass spectrometry.

Pulse-chase experiments and protein turnover. Pulse-chase experiments were done to measure protein half-lives (Materials and Methods). Cells were labeled with [^{35}S]methionine for 10 min, and then an excess of unlabeled methionine was added. Samples were taken at 0, 10, 20, 30, 60, and 90 min after the beginning of the chase. Equal amounts of ^{35}S were loaded from each sample; 2D gels were run, and spots were quantitated. Surprisingly, almost every spot was nearly constant in amount of radioactivity over the entire time course (not shown). A few spots shifted from one position to another because of post-translational modifications (e.g., phosphorylation of Rpa0 and Efb1). Thus, the proteins being visualized are all or nearly all very stable proteins, with half-lives of more than 90 min. Gygi et al. (10) have come to a similar conclusion by using the N-end rule to predict protein half-lives. This result does not imply that all yeast proteins are stable. The proteins being visualized are abundant proteins; this is partly because they are stable proteins.

Protein quantitation. Because all of the proteins seen had effectively the same half-life, the abundance of each protein was directly proportional to the amount of radioactivity incorporated during labeling. Thus, after taking into account the total number of protein molecules per cell, the average content of methionine and cysteine, and the methionine and cysteine content of each identified protein, we could calculate the abundance of each identified protein (Tables 1 and 2; Materials and Methods). About 1,000 unidentified proteins were also quantified, assuming an average content of Met and Cys.

Many proteins give multiple spots (7, 8). The contribution from each spot was summed to give the total protein amount. However, many proteins probably have minor spots that we are not aware of, causing the amount of protein to be underestimated.

When the proteins on a pH 4–8 gel were ordered by abundance, the most abundant protein had 8,904 ppm, the 10th most abundant had 2,842 ppm, the 100th most abundant had 314 ppm, the 500th most abundant had 57 ppm, and the 1,000th most abundant (visualized at greater than optimum exposure) had 23 ppm. Thus, there is more than a 300-fold range in abundance among the visualized proteins. The most abundant 10 proteins account for about 25% of the total protein on the pH 4–8 gel, the most abundant 60 proteins account for 50%, and the most abundant 500 proteins account for 80%. Since it seems likely that the pH 4–8 gels give a representative sampling of all proteins, we estimate that half of the total cellular protein is accounted for by fewer than 100 different gene products, principally glycolytic enzymes and proteins involved in protein synthesis.

Correlation of protein abundance with mRNA abundance. Estimates of mRNA abundance for each gene have been made by SAGE (27) and by hybridization of cRNA to oligonucleotide arrays (30). These two methods give broadly similar results, yet each method has strengths and weaknesses (Materials and Methods). Table 1 lists the number of molecules of mRNA per cell for each gene studied. One measurement (mRNA) uses data from SAGE analysis alone (27); a second incorporates data from both SAGE and hybridization (30) (adjusted mRNA) (Table 1; Materials and Methods). We correlated protein abundance with mRNA abundance (Fig. 2). For adjusted mRNA versus protein, the Spearman rank correlation coefficient, r_s, was 0.74 ($P < 0.0001$), and the Pearson correlation coefficient, r_p, on log transformed data (Materials and Methods) was 0.76 ($P < 0.00001$). We obtained similar correlations for mRNA versus protein and also for other data transformations (Materials and Methods). Thus, several statistical methods show a strong and significant correlation between mRNA abundance and protein abundance. Of course, the correlation is far from perfect; for mRNAs of a given abundance, there is at least a 10-fold range of protein abundance (Fig. 2). Some of this scatter is probably due to posttranscriptional regulation, and some is due to errors in the mRNA or protein data. For example, the protein Yef3 runs poorly on our gels, giving multiple smeared spots. Its abundance has probably been underestimated, partly explaining the low protein/mRNA ratio of Yef3. It is the most extreme outlier in Fig. 2.

These data on mRNA (27, 30) and protein abundance (Table 1) suggest that for each mRNA molecule, there are on average 4,000 molecules of the cognate protein. For instance, for Act1 (actin) there are about 54 molecules of mRNA per cell and about 205,000 molecules of protein. Assuming an mRNA half-life of 30 min (12) and a cell doubling time of 120 min, this suggests that an individual molecule of mRNA might be translated roughly 1,000 times. These calculations are limited to mRNAs for abundant proteins, which are likely to be the mRNAs that are translated best.

A full complement of cell protein is synthesized in about 120 min under these conditions. Thus, 4,000 molecules of protein per molecule of mRNA implies that translation initiates on an mRNA about once every 2 s. This is a remarkably high rate; it implies that if an average mRNA bears 10 ribosomes engaged in translation, then each ribosome completes translation in 20 s; if an average protein has 450 residues; this in turn implies translation of over 20 amino acids per s, a rate considerably higher than estimated for mammalians (3 to 8 amino acids per s) (18). These estimates depend on the amount of mRNA per cell (11, 27).

The large number of protein molecules that can be made from a single mRNA raises the issue of how abundance is controlled for less abundant proteins. Many nonabundant proteins may be unstable, and this would reduce the protein/mRNA ratio. In addition, many nonabundant proteins may be translated at suboptimal rates. We have found that mRNAs for nonabundant proteins usually have suboptimal contexts for translational initiation. For example, there are over 600 yeast genes which probably have short open reading frames in the mRNA upstream of the main open reading frame (17a). These may be devices for reducing the amount of protein made from a molecule of mRNA.

Correlation of codon bias with protein abundance. The mRNAs for highly expressed proteins preferentially use some codons rather than others specifying the same amino acid (14). This preference is called codon bias. The codons preferred are those for which the tRNAs are present in the greatest amounts. Use of these codons may make translation faster or more efficient and may decrease misincorporation. These effects are most important for the cell for abundant proteins, and so codon bias is most extreme for abundant proteins. The effect can be dramatic—highly biased mRNAs may use only 25 of the 61 codons.

We asked whether the correlation of codon bias with abundance continues for medium-abundance proteins. There are various mathematical expressions quantifying codon bias; here, we have used the CAI (24) (Materials and Methods) because it gives a result between 0 and 1. The r_s for CAI versus protein abundance is 0.80 ($P < 0.0001$), similar to the mRNA-protein

TABLE 1. Quantitative data[a]

Function	Name	CAI	mRNA	Adjusted mRNA	Protein (Glu) (10^3)	Protein (Eth) (10^3)	E/G ratio
Carbohydrate metabolism	Adh1	0.810	197	197	1,230	972	0.79
	Adh2	0.504	0		0	963	>20
	Cit2	0.185	1	2.8	23	288	12
	Eno1	0.870	No *Nla*		410	974	2.4
	Eno2	0.892	248	248	650	215	0.33
	Fba1	0.868	179	179	640	608	0.95
	Hxk1,2	0.500	13	10.5	62	46	
	Icl1	0.251	0		0	671	>20
	Pdb1	0.342	5	5	41	33	
	Pdc1	0.903	226	226	280	205	0.73
	Pfk1	0.465	5	5	75	53	0.71
	Pgi1	0.681	14	14	160	120	0.75
	Pyc1	0.260	1	0.7	37	34	
	Tal1	0.579	5	5	110	35	
	Tdh2	0.904	63	63	430	876	NR
	Tdh3	0.924	460	460	1,670	1,927	NR
	Tpi1	0.817	No *Nla*		No Met	No Met	
Protein synthesis	Efb1	0.762	33	16.5	358	362	
	Eft1,2	0.801	26	26	99	54	0.55
	Prt1	0.303	4	0.7	12	6	
	Rpa0	0.793	246	246	277	100	0.36
	Tif1,2	0.752	29	29	233	106	0.46
	Yef3	0.777	36	36	14	ND	
Heat shock	Hsc82	0.581	2	2.9	112	75	0.67
	Hsp60	0.381	9	2.3	35	82	2.3
	Hsp82	0.517	2	1.3	52	135	2.6
	Hsp104	0.304	7	7	70	161	2.3
	Kar2	0.439	5	10.1	43	102	2.4
	Ssa1	0.709	2	4.3	303	421	1.4
	Ssa2	0.802	10	5	213	324	1.5
	Ssb1,2	0.850	50	50	270	85	
	Ssc1	0.521	2	2.6	68	80	1.2
	Sse1	0.521	8	8	96	48	
	Sti1	0.247	1	1.1	25	44	1.7
Amino acid synthesis	Ade1	0.229	4	4	14	27	
	Ade3	0.276	2	1.7	12	9	
	Ade5,7	0.257	2	1.4	14	4	
	Arg4	0.229	1	8.1	41	41	
	Gdh1	0.585	10	27	148	55	
	Gln1	0.524	11	11	77	104	1.3
	His4	0.267	3	3	15	23	1.5
	Ilv5	0.801	6	6	152	109	0.7
	Lys9	0.332	4	4	32	17	0.52
	Met6	0.657	No *Nla*	22	190	80	0.42
	Pro2	0.248	3	3	30	12	
	Ser1	0.258	2	1.2	15	8	
	Trp5	0.319	5	5	28	12	
Miscellaneous	Act1	0.710	54	54	205	164	0.78
	Adk1	0.531	No *Nla*		47	43	
	Ald6	0.520	3	3	181	159	
	Atp2	0.424	1	4.1	76	109	1.4
	Bmh1	0.322	46	46	191	137	0.72
	Bmh2	0.384	1	1.4	134	147	
	Cdc48	0.306	2	2.4	32	26	
	Cdc60	0.299	2	0.86	6	2	
	Erg20	0.373	5	5	92	39	
	Gpp1	0.603	16	5	234	158	
	Gsp1	0.621	3	3	115	39	0.34
	Ipp1	0.620	4	4	254	147	0.58
	Lcb1	0.173	0.3	0.8	19	40	
	Mol1	0.423	0	0.45	20	16	
	Pab1	0.488	3	3	41	19	0.47
	Psa1	0.600	15	15	148	56	
	Rnr4	0.497	6	6	44	37	
	Sam1	0.494	5	5	59	21	
	Sam2	0.497	3	15	63	20	
	Sod1	0.376	36	36	631	618	
	Uba1	0.212	2	2	14	20	
	YKL056	0.731	62	62	253	112	0.44
	YLR109	0.549	21	21	930		
	YMR116	0.777	41	41	184	40	0.20

[a] CAI, a measure of codon bias, is taken from the YPD. mRNA, number of mRNA molecules per cell from SAGE data (27); adjusted mRNA, number of mRNA molecules per cell based on both SAGE and chip hybridization (30) (see Materials and Methods); Protein (Glu), number of molecules of protein per cell in YNB-glucose; Protein (Eth), number of molecules of protein per cell in YNB-ethanol; E/G ratio, ratio of protein abundance in ethanol to glucose. The E/G ratio is not given if it was close to 1 or if it was not repeatable (NR) in multiple gels. Some gene products (e.g., Tif1 and Tif2 [Tif1,2]) were difficult to distinguish on either a protein or an mRNA basis; these are pooled. No *Nla*, there was no suitable *Nla*III site in the 3′ region of the gene, and so there are no SAGE mRNA data; No Met, the mature gene product contains no methionines, and so there are no reliable protein data.

TABLE 2. Functions of proteins listed in Table 1

Name[a]	YPD title lines[b]
Adh1	Alcohol dehydrogenase I; cytoplasmic isozyme reducing acetaldehyde to ethanol, regenerating NAD^+
Adh2	Alcohol dehydrogenase II; oxidizes ethanol to acetaldehyde, glucose repressed
Cit2	Citrate synthase, peroxisomal (nonmitochondrial); converts acetyl-CoA and oxaloacetate to citrate plus CoA
Eno1	Enolase 1 (2-phosphoglycerate dehydratase); converts 2-phospho-D-glycerate to phosphoenolpyruvate in glycolysis
Eno2	Enolase 2 (2-phosphoglycerate dehydratase); converts 2-phospho-D-glycerate to phosphoenolpyruvate in glycolysis
Fba1	Fructose bisphosphate aldolase II; sixth step in glycolysis
Hxk1	Hexokinase I; converts hexoses to hexose phosphates in glycolysis; repressed by glucose
Hxk2	Hexokinase II; converts hexoses to hexose phosphates in glycolysis and plays a regulatory role in glucose repression
Icl1	Isocitrate lyase, peroxisomal; carries out part of the glyoxylate cycle; required for gluconeogenesis
Pdb1	Pyruvate dehydrogenase complex, E1 beta subunit
Pdc1	Pyruvate decarboxylase isozyme 1
Pfk1	Phosphofructokinase alpha subunit, part of a complex with Pfk2p which carries out a key regulatory step in glycolysis
Pgi1	Glucose-6-phosphate isomerase, converts glucose-6-phosphate to fructose-6-phosphate
Pyc1	Pyruvate carboxylase 1; converts pyruvate to oxaloacetate for gluconeogenesis
Tal1	Transaldolase; component of nonoxidative part of pentose phosphate pathway
Tdh2	Glyceraldehyde-3-phosphate dehydrogenase 2; converts D-glyceraldehyde 3-phosphate to 1,3-dephosphoglycerate
Tdh3	Glyceraldehyde-3-phosphate dehydrogenase 3; converts D-glyceraldehyde 3-phosphate to 1,3-dephosphoglycerate
Tpi1	Triosephosphate isomerase; interconverts glyceraldehyde-3-phosphate and dihydroxyacetone phosphate
Efb1	Translation elongation factor EF-1β; GDP/GTP exchange factor for Tef1p/Tef2p
Eft1	Translation elongation factor EF-2; contains diphthamide which is not essential for activity; identical to Eft2p
Eft2	Translation elongation factor EF-2; contains diphthamide which is not essential for activity; identical to Eft1p
Prt1	Translation initiation factor eIF3 beta subunit (p90); has an RNA recognition domain
Rpa0 (RPPO)	Acidic ribosomal protein A0
Tif1	Translation initiation factor 4A (eIF4A) of the DEAD box family
Tif2	Translation initiation factor 4A (eIF4A) of the DEAD box family
Yef3	Translation elongation factor EF-3A; member of ATP-binding cassette superfamily
Hsc82	Chaperonin homologous to *E. coli* HtpG and mammalian HSP90
Hsp60	Mitochondrial chaperonin that cooperates with Hsp10p; homolog of *E. coli* GroEL
Hsp82	Heat-inducible chaperonin homologous to *E. coli* HtpG and mammalian HSP90
Hsp104	Heat shock protein required for induced thermotolerance and for resolubilizing aggregates of denatured proteins; important for [psi$^-$]-to-[PSI$^+$] prion conversion
Kar2	Heat shock protein of the endoplasmic reticulum lumen required for protein translocation across the endoplasmic reticulum membrane and for nuclear fusion; member of the HSP70 family
Ssa1	Cytoplasmic chaperone; heat shock protein of the HSP70 family
Ssa2	Cytoplasmic chaperone; member of the HSP70 family
Ssb1	Heat shock protein of HSP70 family involved in the translational apparatus
Ssb2	Heat shock protein of HSP70 family, cytoplasmic
Ssc1	Mitochondrial protein that acts as an import motor with Tim44p and plays a chaperonin role in receiving and folding of protein chains during import; heat shock protein of HSP70 family
Sse1	Heat shock protein of the HSP70 family; multicopy suppressor of mutants with hyperactivated Ras/cyclic AMP pathway
Sti1	Stress-induced protein required for optimal growth at high and low temperature; has tetratricopeptide repeats
Ade1	Phosphoribosylamidoimidazole-succinocarboxamide synthase: catalyzes the seventh step in de novo purine biosynthesis pathway
Ade3	C_1 tetrahydrofolate synthase (trifunctional enzyme), cytoplasmic
Ade5,7	Phosphoribosylamine-glycine ligase plus phosphoribosylformylglycinamidine cyclo-ligase; bifunctional protein
Arg4	Argininosuccinate lyase; catalyzes the final step in arginine biosynthesis
Gdh1	Glutamate dehydrogenase ($NADP^+$); combines ammonia and α-ketoglutarate to form glutamate
Gln1	Glutamine synthetase; combines ammonia to glutamate in ATP-driven reaction
His4	Phosphoribosyl-AMP cyclohydrolase/phosphoribosyl-ATP pyrophosphohydrolase/histidinol dehydrogenase; 2nd, 3rd, and 10th steps of his biosynthesis pathway
Ilv5	Ketol-acid reductoisomerase (acetohydroxy, acid reductoisomerase) (alpha-keto-β-hydroxylacyl) reductoisomerase); second step in Val and Ilv biosynthesis pathway
Lys9	Saccharopine dehydrogenase ($NADP^+$, L-glutamate forming) (saccharopine reductase), seventh step in lysine biosynthesis pathway
Met6	Homocysteine methyltransferase; (5-methyltetrahydropteroyl triglutamate-homocysteine methyltransferase), methionine synthase, cobalamin independent
Pro2	γ-Glutamyl phosphate reductase (phosphoglutamate dehydrogenase), proline biosynthetic enzyme
Ser1	Phosphoserine transaminase; involved in synthesis of serine from 3-phosphoglycerate
Trp5	Tryptophan synthase, last (5th) step in tryptophan biosynthesis pathway
Act1	Actin; involved in cell polarization, endocytosis, and other cytoskeletal functions
Adk1	Adenylate kinase (GTP:AMP phosphotransferase), cytoplasmic
Ald6	Cytosolic acetaldehyde dehydrogenase
Atp2	Beta subunit of F1-ATP synthase; 3 copies are found in each F1 oligomer
Bmh1	Homolog of mammalian 14-3-3 protein; has strong similarity to Bmh2p
Bmh2	Homolog of mammalian 14-3-3 protein; has strong similarity to Bmh1p
Cdc48	Protein of the AAA family of ATPases; required for cell division and homotypic membrane fusion
Cdc60	Leucyl-tRNA synthetase, cytoplasmic
Erg20	Farnesyl pyrophosphate synthetase; may be rate-limiting step in sterol biosynthesis pathway
Gpp1 (Rhr2)	DL-Glycerol phosphate phosphatase
Gsp1	Ran, a GTP-binding protein of the Ras superfamily involved in trafficking through nuclear pores
Ipp1	Inorganic pyrophosphatase, cytoplasmic
Lcb1	Component of serine C-palmitoyltransferase; first step in biosynthesis of long-chain base component of sphingolipids
Mol1 (Thi4)	Thiamine-repressed protein essential for growth in the absence of thiamine
Pab1	Poly(A)-binding protein of cytoplasm and nucleus; part of the 3′-end RNA-processing complex (cleavage factor I); has 4 RNA recognition domains
Psa1	Mannose-1-phosphate guanyltransferase; GDP-mannose pyrophosphorylase
Rnr4	Ribonucleotide reductase small subunit
Sam1	*S*-Adenosylmethionine synthetase 1
Sam2	*S*-Adenosylmethionine synthetase 2
Sod1	Copper-zinc superoxide dismutase
Uba1	Ubiquitin-activating (E1) enzyme
YKL056	Resembles translationally controlled tumor protein of animal cells and higher plants
YLR109 (Ahp1)	Alkyl hydroperoxide reductase
YMR116 (Asc1)	Abundant protein with effects on translational efficiency and cell size, has two WD (WD-40) repeats

[a] Accepted name from the *Saccharomyces* genome database and YPD. Names in parentheses represent recent changes.
[b] Courtesy of Proteome, Inc., reprinted with permission.

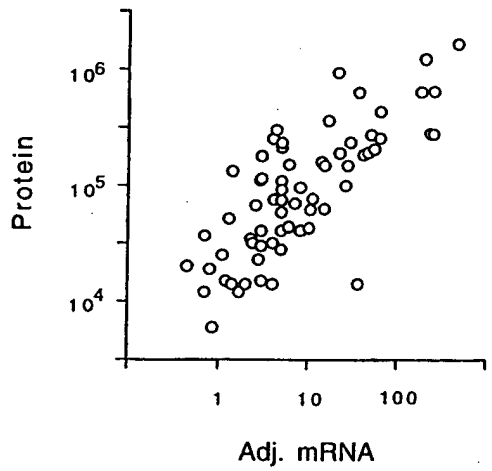

FIG. 2. Correlation of protein abundance with adjusted mRNA abundance. The number of molecules per cell of each protein is plotted against the number of molecules per cell of the cognate mRNA, with an r_p of 0.76. Note the logarithmic axes. Data for mRNA were taken from references 27 and 30 and combined as described in Materials and Methods.

correlation, confirming a strong correlation between CAI and protein abundance (Fig. 3). The relationship between CAI and protein abundance is log linear from about 1,000,000 to about 10,000 molecules per cell. We have no data for rarer proteins.

It is not clear whether CAI reflects maximum or average levels of protein expression. The proteins used for the CAI-protein correlation included some proteins which were not expressed at maximum levels under the condition of the experiment (Hsc82, Hsp104, Ssa1, Ade1, Arg4, His4, and others). When these proteins were removed from consideration and the correlation between CAI and the remaining (presumably constitutive) proteins was recalculated, the r_s was essentially unchanged (not shown).

The equation describing the graph in Fig. 3 is log (protein molecules/cell) = (2.3 × CAI) + 3.7. Thus, under certain conditions (a CAI of 0.3 or greater; a constitutively expressed gene), a very rough estimate of protein abundance can be made by raising 10 to the power of [(2.3 × CAI) + 3.7].

The distribution of CAI over the genome (Fig. 4) consists of a lower, bell-shaped distribution, possibly indicating a region where there is no selection for codon bias, and an upper, flat distribution, starting at a CAI of about 0.3, possibly indicating a region where there is selection for codon bias. Almost all of the proteins whose abundance we have measured are in the upper, flat portion of the distribution. In the lower, bell-shaped region, we do not know whether there is a correlation between CAI and protein abundance.

Changes in protein abundance in glucose and ethanol. A comparison of cells grown in glucose (Fig. 1A) with cells grown in ethanol (Fig. 1B) is shown in Table 1. As is well known, some proteins are induced tremendously during growth on ethanol. Two striking examples are the peroxisomal enzymes Icl1 (isocitrate lyase) and Cit2 (citrate synthase), which are induced in ethanol by more than 100- and 12-fold, respectively (Fig. 1; Table 1). These enzymes are key components of the glyoxylate shunt, which diverts some acetyl coenzyme A (acetyl-CoA) from the tricarboxylic acid cycle to gluconeogenesis. *S. cerevisiae* requires large amounts of carbohydrate for its cell wall; in ethanol medium, this carbohydrate comes from gluconeogenesis, which depends on the glyoxylate shunt and on the glycolytic pathway running in reverse. The need for gluconeogenesis also explains why glycolytic enzymes are abundant even in ethanol medium. Thus, 2D gel analysis shows the prominence of the glycolytic and glyoxylate shunt enzymes in cells grown on ethanol, emphasizing that gluconeogenesis, presumably largely for production of the cell wall, is a major metabolic activity under these conditions.

During gluconeogenesis, substrate-product relationships are reversed for the glycolytic enzymes. One might expect that not all glycolytic enzymes would be well adapted to the reverse reaction. Indeed, 2D gels show that in ethanol, Adh2 (alcohol dehydrogenase 2) is strongly induced (16), while its isozyme Adh1 is not greatly affected. Adh1 and Adh2 each interconvert acetaldehyde and ethanol. Adh1 has a relatively high K_m for ethanol (17 mM), while Adh2 has a lower K_m (0.8 mM) (5). Thus, it is thought that Adh1 is specialized for glycolysis (acetaldehyde to ethanol), while Adh2 is specialized for respiration (ethanol to acetaldehyde) (5, 29). Similarly, Eno1 (enolase 1) is induced in ethanol, while its isozyme Eno2 (enolase 2) decreases in abundance (Table 1) (4, 19). Eno1 is inhibited by 2-phosphoglycerate (the glycolytic substrate), while Eno2 is inhibited by phosphoenolpyruvate (the gluconeogenic substrate) (4). Perhaps Eno1 has a lower K_m for phosphoenolpyruvate than does Eno2, though to our knowledge this has not been tested. Thus, the 2D gels distinguish isozymes specialized for growth on glucose (Adh1 and Eno2) from isozymes specialized for ethanol (Adh2 and Eno1).

Many heat shock proteins (e.g., Hsp60, Hsp82, Hsp104, and Kar2) were about twofold more abundant in ethanol medium than in glucose medium. This is consistent with the increased heat resistance of cells grown in ethanol (3).

Enzymes involved in protein synthesis (Eft1, Rpa0, and Tif1) were about twice as abundant in glucose medium as in ethanol medium. This may reflect the higher growth rate of the cells in glucose.

Phosphorylation of proteins. To examine protein phosphorylation, we labeled cells with ^{32}P and ran 2D gels to examine phosphoproteins. About 300 distinct spots, probably representing 150 to 200 proteins, could be seen on pH 4–8 gels (Fig. 5B). We then aligned autoradiograms of three gels, each with a different kind of labeled protein (^{32}P only [Fig. 5B], ^{32}P plus ^{35}S [Fig. 5A], and ^{35}S only [not shown, but see Fig. 1 for example]). In this way, we made provisional identification of

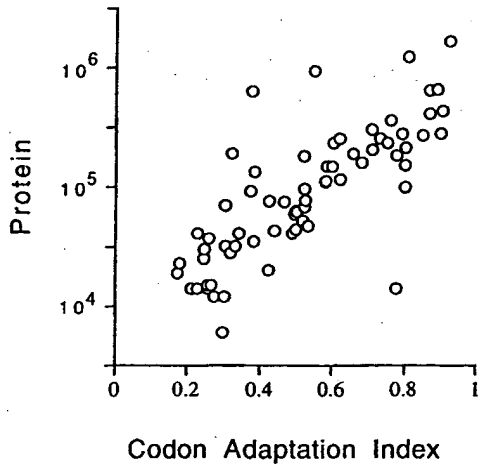

FIG. 3. Correlation of protein abundance with CAI. The number of molecules per cell of each protein is plotted against the CAI for that protein. Note the logarithmic scale on the protein axis. Data for the CAI are from the YPD database (13).

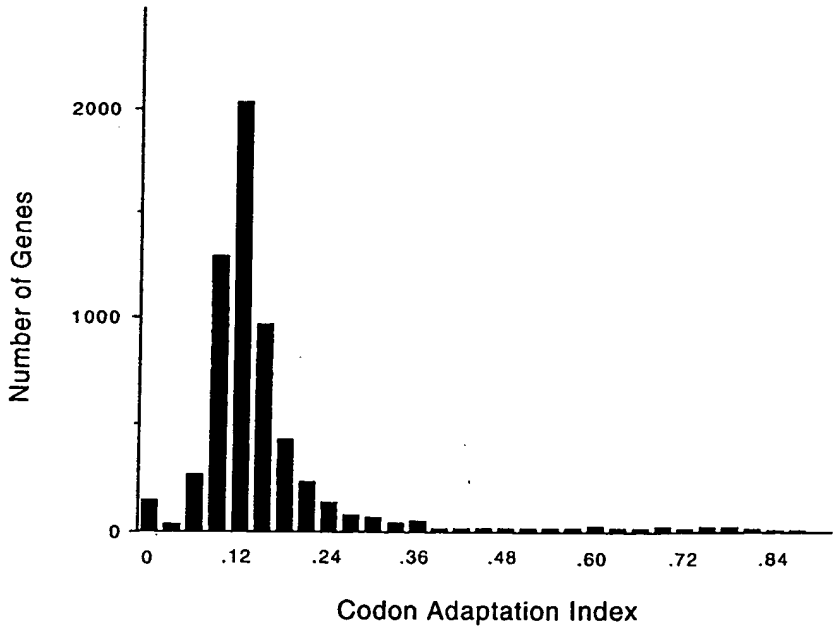

FIG. 4. Distribution of CAI over the whole genome, shown in intervals of 0.030 (i.e., there are 150 genes with a CAI between 0.000 and 0.030, inclusive; 31 genes with a CAI between 0.031 and 0.060; 269 genes with a CAI between 0.061 and 0.090; 1,296 genes with a CAI between 0.091 and 0.120; etc.). The distribution peaks with 2,028 genes with a CAI between 0.121 and 0.150.

some of the ^{32}P-labeled spots as particular ^{35}S-labeled spots. All such identifications are somewhat uncertain, since precise alignments are difficult, and of course multiple spots may exactly comigrate. Nevertheless, we believe that most of the provisional identifications are probably correct. Among the major ^{32}P-labeled proteins are the hexokinases Hxk1 and Hxk2, the acidic ribosome-associated protein Rpa0, the translation factors Yef3 and Efb1, and probably Hsp70 heat shock proteins of the Ssa and Ssb families. Rpa0 and Efb1 are quantitatively monophosphorylated.

Many yeast proteins resolve into multiple spots on these 2D gels (7). Yef3 has five or more spots, at least four of which comigrate with ^{32}P. Tpi1 has a major spot showing no ^{32}P labeling and a minor, more acidic spot which overlaps with some ^{32}P label. Tif1 has at least seven spots (7); two of these overlap with some ^{32}P label, but five do not (Fig. 5). Eft1 has at least three spots (7), and none of these overlap with ^{32}P, although there are three nearby, unidentified ^{32}P-labeled spots (a, c, and d in Fig. 5). Spots that seem to be extra forms of Met6, Pdc1, Eno2, and Fba1 can be seen in Fig. 6A, but there is little ^{32}P at these positions in Fig. 5. Thus, phosphorylation explains some but not all of the different protein isoforms seen.

The cell cycle is regulated in part by phosphorylation. We compared ^{32}P-labeled proteins from cells synchronized in G_1 with α-factor, in cells synchronized in G_1 by depletion of G_1 cyclins, and in cells synchronized in M phase with nocodazole. Only very minor differences were seen, and these were difficult to reproduce. The cell cycle proteins regulated by phosphorylation may not be abundant enough for this technique to be applied easily.

Centrifugal fractionation. We fractionated ^{35}S-labeled extracts by centrifugation (Materials and Methods). Figure 6A shows the proteins in the supernatant of a high-speed (100,000 × *g*, 30 min) centrifugation, while Fig. 6B shows the proteins in the pellet of a low-speed (16,000 × *g*, 10 min) centrifugation. Many proteins are tremendously enriched in one fraction or the other, while others are present in both. Most glycolytic enzymes (e.g., Tdh2, Tdh3, Eno2, Pdc1, Adh1, and Fba1) are enriched in the supernatant fraction. The only exception is Pfk1 (not indicated), which is found in both pellet and supernatant fractions. Many proteins involved in protein synthesis (Eft1, Yef3, Prt1, Tif1, and Rpa0) are in the pellet, possibly because of the association of ribosomes with the endoplasmic reticulum. However, Efb1 is in the supernatant, as is a substantial portion of the Eft1. Perhaps surprisingly, several mitochondrial proteins (Atp2 [not shown] and Ilv5) are largely in the supernatant. Perhaps glass bead breakage of cells releases mitochondrial proteins. The nuclear protein Gsp1 is in the pellet fraction. The enrichment produced by centrifugation makes it possible to see minor spots which are otherwise poorly resolved from surrounding proteins. Figure 6B shows that the previously identified Tif1 spot is surrounded by as many as six other spots that cofractionate. We observed six identical or very similar additional spots when we overexpressed Tif1 from a high-copy-number plasmid (not shown). Signal overlaps only one or two of these spots in ^{32}P-labeling experiments (Fig. 5), and so the different forms are not mainly due to different phosphorylation states.

DISCUSSION

Our experience with developing a 2D gel protein database for *S. cerevisiae* is summarized here. With current technology, we can see the most abundant 1,200 proteins, which is about one-third to one-quarter of the proteins expressed. The remaining proteins will be difficult to see and study with the methods that we have used, not because of a lack of sensitivity but because weak spots are covered by nearby strong spots.

Of the 1,200 proteins seen, we have identified 148, with a bias toward the most abundant proteins. Steady application of the methods already used would allow identification of most of the remaining proteins. Gene overexpression will be particularly useful, since it is not affected by the lower abundance of the remaining visible proteins.

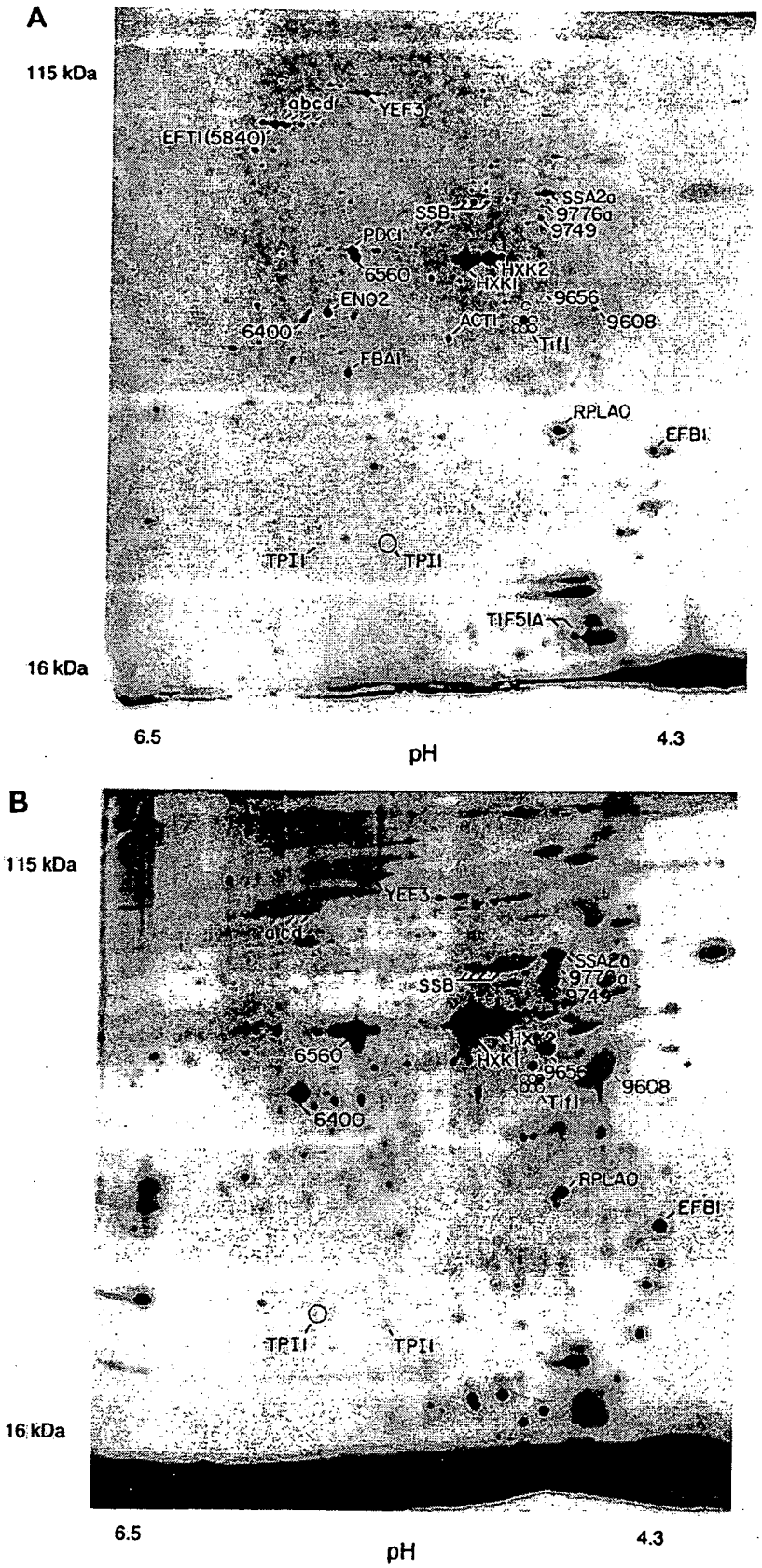

FIG. 5. Phosphorylated proteins. (A) Mixture of ^{32}P-labeled proteins and ^{35}S-labeled proteins. Two separate labeling reactions were done, one with ^{32}P and one with ^{35}S, and extracts were mixed and run on a 2D gel. Spots marked with numbers rather than gene names represent spots noted on ^{35}S gels but unidentified. Spots labeling with ^{32}P were identified by (i) increased labeling compared to the ^{35}S-only gel (not shown); (ii) the characteristic fuzziness of a ^{32}P-labeled spot; and (iii) the decay of signal intensity seen on exposures made 4 weeks later (not shown). A minor form of Tpi1 and at least six minor forms of Tif1 have been noted in overexpression experiments (see also Fig. 6B); positions of the minor forms are indicated by circles. (B) ^{32}P-only labeling. The major form of Tpi1, which is not labeled with ^{32}P, is indicated by a large circle; positions of seven forms of Tif1 are indicated by smaller circles.

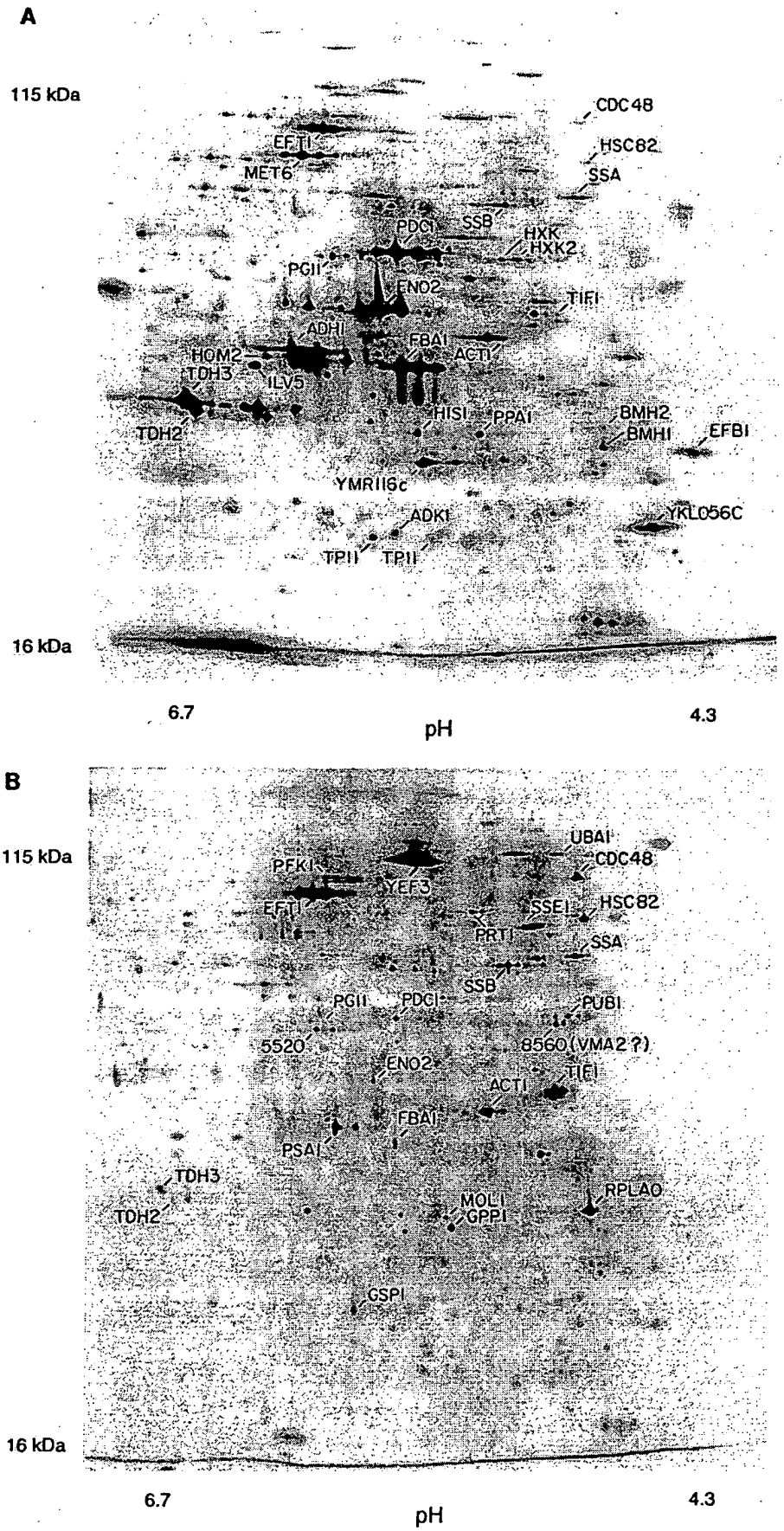

FIG. 6. Fractionation by centrifugation. (A) Proteins in the supernatant of a 100,000 × *g*, 30-min spin; proteins in the pellet of a 16,000 × *g*, 10-min spin. Supernatant fractions examined in multiple experiments done over a wide range of *g* forces looked similar to each other, as did the pellet fractions.

2D gels of the kind that we have used are not suitable for visualization of rare proteins. However it will be possible to study on a global basis metabolic processes involving relatively abundant proteins, such as protein synthesis, glycolysis, gluconeogenesis, amino acid synthesis, cell wall synthesis, nucleotide synthesis, lipid metabolism, and the heat shock response.

Gygi et al. (10) have recently completed a study similar to ours. Despite generating broadly similar data, Gygi et al. reached markedly different conclusions. We believe that both mRNA abundance and codon bias are useful predictors of protein abundance. However, Gygi et al. feel that mRNA abundance is a poor predictor of protein abundance and that "codon bias is not a predictor of either protein or mRNA levels" (10). These different conclusions are partly a matter of viewpoint. Gygi et al. focus on the fact that the correlations of mRNA and codon bias with protein abundance are far from perfect, while we focus on the fact that, considering the wide range of mRNA and protein abundance and the undoubted presence of other mechanisms affecting protein abundance, the correlations are quite good.

However, the different conclusions are also partly due to different methods of statistical analysis and to real differences in data. With respect to statistics, Gygi et al. used the Pearson product-moment correlation coefficient (r_p) to measure the covariance of mRNA and protein abundance. Depending on the subset of data included, their r_p values ranged from 0.1 to 0.94. Because of the low r_p values with some subsets of the data, Gygi et al. concluded that the correlation of mRNA to protein was poor. However, the r_p correlation is a parametric statistic and so requires variates following a bivariate normal distribution; that is, it would be valid only if both mRNA and protein abundances were normally distributed. In fact, both distributions are very far from normal (data not shown), and so a calculation of r_p is inappropriate. There was no statistical backing for the assertion that codon bias fails to predict protein abundance.

We have taken two statistical approaches. First, we have used the Spearman rank correlation coefficient (r_s). Since this statistic is nonparametric, there is no requirement for the data to be normally distributed. Using the r_s, we find that mRNA abundance is well correlated with protein abundance (r_s = 0.74), and the CAI is also well correlated with protein abundance (r_s = 0.80) (and also with mRNA abundance [data not shown]). For the data of Gygi et al. (10), we obtained similar results, though with their data the correlation is not as good; r_s = 0.59 for the mRNA-to-protein correlation, and r_s = 0.59 for the codon bias-to-protein correlation.

In a second approach, we transformed the mRNA and protein data to forms where they were normally distributed, to allow calculation of an r_p (Materials and Methods). Two transformations, Box-Cox and logarithmic, were used; both gave good correlations with our data [e.g., r_p = 0.76 for log(adjusted RNA) to log(protein)]. We were not able to transform the data of Gygi et al. to a normal distribution.

Finally, there are also some differences in data between the two studies. These may be partly due to the different measurement techniques used: Gygi et al. measured protein abundance by cutting spots out of gels and measuring the radioactivity in each spot by scintillation counting, whereas we used phosphorimaging of intact gels coupled to image analysis. We compared our data to theirs for the proteins common between the studies (but excluding proteins whose mRNAs are known to differ between rich and minimal media, and excluding Tif1, which was anomalous in differing by 100-fold between the two data sets). The r_s between the two protein data sets was 0.88 ($P < 0.0001$). Although this is a strong correlation, the fact that it is less than 1.0 suggests that there may have been errors in measuring protein abundance in one or both studies. After normalizing the two data sets to assume the same amount of protein per cell, we found a systematic tendency for the protein abundance data of Gygi et al. to be slightly higher than ours for the highest-abundance proteins and also for the lowest-abundance proteins but slightly lower than ours for the middle-abundance proteins. These systematic differences suggest some systematic errors in protein measurement. Although we do not know what the errors are, we suggest the following as a reasonable speculation. For the highest-abundance proteins, we may have underestimated the amount of protein because of a slightly nonlinear response of the phosphorimager screens. For the lowest-abundance proteins, Gygi et al. may have overestimated the amount of protein because of difficulties in accurately cutting very small spots out of the gel and because of difficulties in background subtraction for these small, weak spots. The difference in the middle abundance proteins may be a consequence of normalization, given the two errors above.

The low-abundance proteins in the data set of Gygi et al. have a poor correlation with mRNA abundance. We calculate that the r_s is 0.74 for the top 54 proteins of Gygi et al. but only 0.22 for the bottom 53 proteins, a statistically significant difference. However, with our data set, the r_s is 0.62 for the top 33 proteins and 0.56 (not significantly different) for the bottom 33 proteins (which are comparable in abundance to the bottom 53 proteins of Gygi et al.). Thus, our data set maintains a good correlation between mRNA and protein abundance even at low protein abundance. This is consistent with our speculation that protein quantification by phosphorimaging and image analysis may be more accurate for small, weak spots than is cutting out spots followed by scintillation counting. Our relatively good correlations even for nonabundant proteins may also reflect the fact that we used both SAGE data and RNA hybridization data, which is most helpful for the least abundant mRNAs. In summary, we feel that the poor correlation of protein to mRNA for the nonabundant proteins of Gygi et al. may reflect difficulty in accurately measuring these nonabundant proteins and mRNAs, rather than indicating a truly poor correlation in vivo. It is not surprising that observed correlations would be poorer with less-abundant proteins and mRNAs, simply because the accuracy of measurement would be worse.

How well can mRNA abundance predict protein abundance? With r_p = 0.76 for logarithmically transformed mRNA and protein data, the coefficient of determination, $(r_p)^2$, is 0.58. This means that more than half (in log space) of the variation in protein abundance is explained by variation in mRNA abundance. When converted back to arithmetic values, protein abundances vary over about 200-fold (Table 1), and $(r_p)^2$ = 0.58 for the log data means that of this 200-fold variation, about 20-fold is explained by variation in the abundance of mRNA and about 10-fold is unexplained (but could be due partly to measurement errors). For proteins much less abundant than those considered here, we imagine the *in vivo* correlation between mRNA and protein abundance will be worse, and other regulatory mechanisms such as protein turnover will be more important.

Some important conclusions can be drawn from this sampling of the proteome. First, there is an enormous range of protein abundance, from nearly 2,000,000 molecules per cell for some glycolytic enzymes to about 100 per cell for some cell cycle proteins (26a). Second, about half of all cellular protein is found in fewer than 100 different gene products, which are mostly involved in carbohydrate metabolism or protein synthe-

sis. Third, the correlation between protein abundance and CAI is log linear as far as we can see, which is from about 10,000 protein molecules per cell to about 1,000,000. This is somewhat surprising, because it implies that selective forces for codon bias are significant even at moderate expression levels. It also means that codon bias is a useful predictor of protein abundance even for moderately low bias proteins. Fourth, there is a good correlation between protein abundance and mRNA abundance for the proteins that we have studied. This validates the use of mRNA abundance as a rough predictor of protein abundance, at least for relatively abundant proteins. Fifth, for these abundant proteins, there are about 4,000 molecules of protein for each molecule of mRNA. This last conclusion raises questions as to how the levels of nonabundant proteins are regulated and suggests that protein instability, regulated translation, suboptimal rates of translation, and other mechanisms in addition to transcriptional control may be very important for these proteins.

ACKNOWLEDGMENTS

We thank Neena Sareen and Nick Bizios (CSHL 2D gel laboratory) for production of 2D gels, Tom Volpe for help with some experiments, Corine Driessens for help with calculations and statistics, and Herman Wijnen and Nick Edgington for comments on the manuscript. We especially thank Tim Tully for in-depth statistical analysis and for insightful discussions on statistical interpretations.

This work was supported by grant P41-RR02188 from the NIH Biomedical Research Technology Program, Division of Research Resources, to J.I.G., by Small Business Innovation Research grant R44 GM54110 to Proteome, Inc., by grant DAMD17-94-J4050 from the Army Breast Cancer Program to B.F., and by NIH grant RO1 GM45410 to B.F.

REFERENCES

1. **Baroni, M. D., E. Martegani, P. Monti, and L. Alberghina.** 1989. Cell size modulation by *CDC25* and *RAS2* genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 9:2715–2723.
2. **Boucherie, H., F. Sagliocco, R. Joubert, I. Maillet, J. Labarre, and M. Perrot.** 1996. Two-dimensional gel protein database of Saccharomyces cerevisiae. Electrophoresis 17:1683–1699.
3. **Elliott, B., and B. Futcher.** 1993. Stress resistance of yeast cells is largely independent of cell cycle phase. Yeast 9:33–42.
4. **Entian, K. D., B. Meurer, H. Kohler, K. H. Mann, and D. Mecke.** 1987. Studies on the regulation of enolases and compartmentation of cytosolic enzymes in Saccharomyces cerevisiae. Biochim. Biophys. Acta 923:214–221.
5. **Ganzhorn, A. J., D. W. Green, A. D. Hershey, R. M. Gould, and B. V. Plapp.** 1987. Kinetic characterization of yeast alcohol dehydrogenases. Amino acid residue 294 and substrate specificity. J. Biol. Chem. 262:3754–3761.
6. **Garrels, J. I.** 1989. The Quest system for quantitative analysis of two-dimensional gels. J. Biol. Chem. 264:5269–5282.
7. **Garrels, J. I., B. Futcher, R. Kobayashi, G. I. Latter, B. Schwender, T. Volpe, J. R. Warner, and C. S. McLaughlin.** 1994. Protein identifications for a Saccharomyces cerevisiae protein database. Electrophoresis 15:1466–1486.
8. **Garrels, J. I., C. S. McLaughlin, J. R. Warner, B. Futcher, G. I. Latter, R. Kobayashi, B. Schwender, T. Volpe, D. S. Anderson, R. Mesquita-Fuentes, and W. E. Payne.** 1997. Proteome studies of S. cerevisiae: identification and characterization of abundant proteins. Electrophoresis 18:1347–1360.
9. **Goffeau, A., B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, and S. G. Oliver.** 1996. Life with 6000 genes. Science 274:563–567.
10. **Gygi, S. P., Y. Rochon, B. R. Franza, and R. Aebersold.** 1999. Correlation between protein and mRNA abundance in yeast. Mol. Cell. Biol. 19:1720–1730.
11. **Hereford, L. M., and M. Rosbash.** 1977. Number and distribution of polyadenylated RNA sequences in yeast. Cell 10:453–462.
12. **Herrick, D., R. Parker, and A. Jacobson.** 1990. Identification and comparison of stable and unstable mRNAs in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 10:2269–2284.
13. **Hodges, P. E., A. H. McKee, B. P. Davis, W. E. Payne, and J. I. Garrels.** 1999. The Yeast Proteome Database (YPD): a model for the organization of genome-wide functional data. Nucleic Acids Res. 27:69–73.
14. **Ikemura, T.** 1985. Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. 2:13–34.
15. **Johnston, G. C., F. R. Pringle, and L. H. Hartwell.** 1977. Coordination of growth with cell division in the yeast S. cerevisiae. Exp. Cell Res. 105:79–98.
16. **Johnston, M., and M. Carlson.** 1992. Regulation of carbon and phosphate utilization, p. 193–281. *In* E. Jones, J. Pringle, and J. Broach (ed.), The molecular and cellular biology of the yeast Saccharomyces. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
17. **Kornblatt, M. J., and A. Klugerman.** 1989. Characterization of the enolase isozymes of rabbit brain: kinetic differences between mammalian and yeast enolases. Biochem. Cell. Biol. 67:103–107.

17a. **Latter, G., and B. Futcher.** Unpublished data.

18. **Mathews, B., N. Sonenberg, and J. W. B. Hershey.** 1996. Origins and targets of translational control, p. 1–29. *In* J. W. B. Hershey, M. B. Mathews, and N. Sonenberg (ed.), Translational control. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
19. **McAlister, L., and M. J. Holland.** 1982. Targeted deletion of a yeast enolase structural gene. Identification and isolation of yeast enolase isozymes. J. Biol. Chem. 257:7181–7188.
20. **Monardo, P. J., T. Boutell, J. I. Garrels, and G. I. Latter.** 1994. A distributed system for two-dimensional gel analysis. Comput. Appl. Biosci. 10:137–143.
21. **O'Farrell, P. H.** 1975. High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. 250:4007–4021.
22. **Patterson, S. D., and G. I. Latter.** 1993. Evaluation of storage phosphor imaging for quantitative analysis of 2-D gels using the Quest II system. BioTechniques 15:1076–1083.
23. **Sagliocco, F., J. C. Guillemot, C. Monribot, J. Capdevielle, M. Perrot, E. Ferran, P. Ferrara, and H. Boucherie.** 1996. Identification of proteins of the yeast protein map using genetically manipulated strains and peptide-mass fingerprinting. Yeast 12:1519–1533.
24. **Sharp, P. M., and W. H. Li.** 1987. The Codon Adaptation Index—a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. 15:281–1295.
25. **Shevchenko, A., O. N. Jensen, A. V. Podtelejnikov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortensen, A. Shevchenko, H. Boucherie, and M. Mann.** 1996. Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc. Natl. Acad. Sci. USA 93:14440–14445.
26. **Thomas, B. J., and R. Rothstein.** 1989. Elevated recombination rates in transcriptionally active DNA. Cell 56:619–630.

26a. **Tyers, M., and B. Futcher.** Unpublished data.

27. **Velculescu, V. E., L. Zhang, W. Zhou, J. Vogelstein, M. A. Basrai, D. E. Bassett, Jr., P. Hieter, B. Vogelstein, and K. W. Kinzler.** 1997. Characterization of the yeast transcriptome. Cell 88:243–251.
28. **Warner, J.** 1991. Labeling of RNA and phosphoproteins in S. cerevisiae. Methods Enzymol. 194:423–428.
29. **Wills, C.** 1976. Production of yeast alcohol dehydrogenase isoenzymes by selection. Nature 261:26–29.

29a. **Wodicka, L.** Personal communication.

29b. **Wodicka, L.** Unpublished data.

30. **Wodicka, L., H. Dong, M. Mittmann, M.-H. Ho, and D. J. Lockhart.** 1997. Genome-wide expression monitoring in Saccharomyces cerevisiae. Nat. Biotechnol. 15:1359–1367.

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☑ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☑ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☑ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.